Multidim Syst Sign Process (2013) 24:503–542
DOI 10.1007/s11045-012-0175-6

A stochastic analysis of distance estimation approaches in single molecule microscopy: quantifying the resolution limits of photon-limited imaging systems

Sripad Ram · E. Sally Ward · Raimund J. Ober

Received: 27 August 2011 / Revised: 12 November 2011 / Accepted: 8 January 2012 /
Published online: 26 January 2012

Abstract Optical microscopy is an invaluable tool to visualize biological processes at the cellular scale. In the recent past, there has been significant interest in studying these processes at the single molecule level. An important question that arises in single molecule experiments concerns the estimation of the distance of separation between two closely spaced molecules. Presently, there exists different experimental approaches to estimate the distance between two single molecules. However, it is not clear as to which of these approaches provides the best accuracy for estimating the distance. Here, we address this problem rigorously by using tools of statistical estimation theory. We derive formulations of the Fisher information matrix for the underlying estimation problem of determining the distance of separation from the acquired data for the different approaches. Through the Cramer-Rao inequality, we derive a lower bound to the accuracy with which the distance of separation can be estimated. We show through Monte-Carlo simulations that the bound can be attained by the maximum likelihood estimator. Our analysis shows that the distance estimation problem is in fact related to the localization accuracy problem, the latter being a distinct problem that deals with how accurately the location of an object can be determined. We have carried out a detailed investigation of the relationship between the Fisher information matrices of the two problems for the different experimental approaches considered here. The paper also addresses the issue of a singular Fisher information matrix, which presents a significant complication when calculating the Cramer-Rao lower bound. Here, we show how experimental design can overcome the singularity. Throughout the paper, we illustrate our results by considering a specific image profile that describe the image of a single molecule.

Keywords Marked point process · Photon statistics · Performance bounds · Fluorescence microscopy · Resolution limits · Rayleigh's criterion

S. Ram · E. S. Ward · R. J. Ober
Department of Immunology, University of Texas Southwestern Medical Center, Dallas, TX, USA

R. J. Ober (✉)
Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX, USA
e-mail: ober@utdallas.edu

1 Introduction

The study of biomolecular interactions that occur within a cell is fundamental to all areas of basic biomedical research. The optical microscope is one of the most preferred tools to study biomolecular interactions, as it enables the direct visualization of these processes in real time. For instance, several technological advances in the past decade have made it possible to image individual biomolecules with an optical microscope even in live biological cells (Moerner 2007; Ober et al. 2004a). In many concrete applications, it is important to know the distance of separation between the biomolecules, as this has significant biological implications. The resolution limit of the optical microscope plays a crucial role in determining the ability to measure the distance of separation between biomolecules. Classical resolution criteria such as Rayleigh's criterion, although extensively used, are well known to be based on heuristic notions that render them inadequate for present day microscopy systems. Therefore quantifying the resolution limit is a very important problem with significant implications on the nature and type of studies that can be carried out with an optical microscope.

Current experimental approaches to studying single molecule interactions can be broadly classified into two categories. In one set of approaches, which we refer to as *the simultaneous detection approach* (Fig. 1a), photon emission from the point sources occurs simultaneously during image acquisition and hence the acquired images contain signal from both point sources (Santos and Young 2000; Ram et al. 2006a; Chao et al. 2009a,b). In the other set of approaches, which we refer to as *the separate detection approach* (Fig. 1c), photon emission from the point sources are temporally separated (e.g. stochastic photoactivation (Betzig et al. 2006; Rust et al. 2006; Hess et al. 2006) and blinking (Lidke et al. 2005; Lagerholm et al. 2006)). Hence the acquired images typically contain signal from only one of the point sources. For both types of approaches, the analysis of the acquired data is carried out using

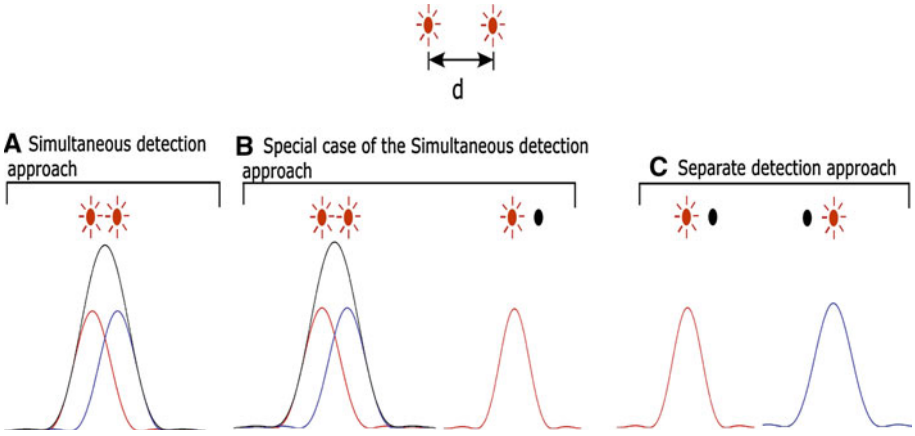

Fig. 1 Different experimental approaches to determine the distance of separation between two identical point sources. **a** illustrates the simultaneous detection approach in which photon emission from both point sources occurs during image acquisition. In this approach, the data consists of a single image that contains signal from both point sources. **b** illustrates the special case of the simultaneous detection approach, where the image of one of the point sources is additionally available. Here, the data consists of a pair of images where one of the images contains signal from only one point source, whereas the other image contains signal from both point sources. **c** illustrates the separate detection approach, where photon emission from the point sources are temporally separated. Here, the data consists of a pair of images, where each image contains signal from either of the point sources

a parameter estimation framework. For example, in the case of the simultaneous detection approach the distance between the point sources is determined by fitting a pair of suitably parameterized image profiles to the acquired data. In the case of the separation detection approach, the analysis involves independently localizing the point sources and then deducing the distance. It has been reported that both approaches are capable of accurately measuring nanometer scale distances, well below the classical resolution criteria. However, an important question arises as to what are the fundamental performance limits of the two experimental approaches to measure the distance of separation.

In this paper, we use the tools of statistical signal processing to investigate this question in a rigorous manner. We formulate the resolution problem as a parameter estimation problem of determining the distance between two closely spaced point sources. The issue of resolvability of the two point sources then becomes a question of how accurately the distance can be estimated, i.e., how large is the standard deviation of the distance estimator. In this context, it is important to know what is the lowest possible standard deviation with which the distance can be estimated, as this can be used as a benchmark for the resolvability of the point sources. For this, we make use of the Cramer-Rao inequality (Rao 1965) which, through the inverse Fisher information matrix, provides a lower bound to the variance of any unbiased estimator of an unknown parameter. Thus, in the present context we interpret the Cramer-Rao lower bound of the distance parameter as a *measure* of resolvability of the two point sources.

Here, we derive formulations of the Fisher information matrix for the parameter estimation problem that underlies the data analysis for the two approaches. Our analysis shows that the Fisher information matrices for the two techniques exhibit very distinct behaviors. For instance, in the simultaneous detection approach the Fisher information matrix depends on the distance of separation between the point sources. In contrast, for the separate detection approach the Fisher information matrix is independent of the distance of separation. As we will see, the distance dependence of the Fisher information matrix has several implications. In particular, for the simultaneous detection approach the Fisher information matrix becomes singular when the distance goes to zero assuming that the two point sources have identical image profiles and photon detection rates, which is typically the case in most imaging applications. An immediate implication is that for very small distances, the Cramer-Rao lower bound of the distance will be numerically very large, thereby predicting poor resolvability of the point sources. On the other hand, the Fisher information matrix for the separate detection approach is invertible for all values of the distance including when the distance is equal to zero.

Another problem that is of significance in the present context is the localization accuracy problem, which deals with how accurately the location of an object can be determined (Wong et al. 2011; Ram et al. 2006b; Ober et al. 2004b; Rohr 2007). For the separate detection approach, the localization accuracy problem naturally arises as part of the data analysis procedure. For the simultaneous detection approach, the localization accuracy problem arises as a special case where in some applications the image of one of the point sources is additionally available (e.g. photobleaching (Ram et al. 2006a; Gordon et al. 2004; Qu et al. 2004), which, in turn, can be used as a priori information (Fig. 1b). Here, we investigate the relationship between the Fisher information matrix of the two approaches and that of the localization accuracy problem. Our analysis shows that for the separate detection approach, the expression for the Fisher information matrix is equivalent to that of the localization accuracy problem, whereas for the simultaneous detection approach the equivalence is attained only when the distance of separation between the point sources becomes very large (i.e., $d \to \infty$). In this context, we also investigate the singularity of the Fisher information matrix for the simultaneous detection approach. In particular we show that the singularity can be removed when the location coordinates of one of the point sources is known a priori.

Previously, we have examined the distance estimation problem for optical microscopes, where we derived analytical expressions for the Fisher information matrix. In Ram et al. (2006a), we investigated the 2D imaging scenario for the simultaneous detection approach, where the point sources were assumed to be located on the x axis of the plane of focus in the object space. In Chao et al. (2009a,b), we considered the 3D imaging scenario for the simultaneous detection approach, where the point sources were assumed to be located anywhere in the object space. In Chao et al. (2009c), we reported numerical calculations of the Cramer-Rao lower bound for the two detection approaches considered here. In the present work, we rigorously analyze the relationship between the Fisher information matrices of the two experimental approaches considered here and that of the localization accuracy problem.

In the past, other groups have investigated the distance estimation problem by adopting a simplified data model, where the acquired data is described as a deterministic signal corrupted by additive noise (Helstrom 1964; Smith 2005; Shahram and Milanfar 2004). Because photon/light emission from a point source is inherently a random phenomenon (Young 1996), it is important to take into account the stochastic nature of the signal (i.e., the photon statistics) from the point sources especially when dealing with photon-limited imaging systems (O'Sullivan et al. 1998). In our (prior and current) work, we have adopted a stochastic framework and model the acquired data as a spatio-temporal random process (marked Poisson process). In this way we explicitly take into account the photon statistics. Thus our results and analyses presented in this paper provide a broad framework to investigate the resolution limits for a wide variety of low light level imaging applications.

The paper is organized as follows. In Sect. 3, we derive general expressions of the Fisher information matrix for the estimation problem that underlies the simultaneous detection approach. We also derive the Fisher information matrix for a concrete scenario in optical microscopy where the image of an object is considered to be spatially invariant. In Sect. 4, we discuss the relationship between the Fisher information matrix for the simultaneous detection approach and that of the localization accuracy problem. In Sect. 5, we consider a special case of the simultaneous detection approach, where we assume that the location coordinates of one of the point sources is known and derive the Fisher information matrix. As we will see, the analysis of this special case provides important insights into the relationship between the two approaches considered here. In Sect. 6, we derive the Fisher information matrix for the separate detection approach. Finally, in Sect. 7 we validate our results by demonstrating that the maximum likelihood estimator of the distance attains the Cramer-Rao lower bound for the different experimental approaches considered here. Throughout the paper, we illustrate our results with examples relevant to single molecule microscopy.

2 Stochastic framework

We assume an acquired image to consist of the time points and the spatial coordinates of the detected photons and model it as a spatio-temporal random process. We refer to this process as the *image detection process* $\mathcal{G}$ (see Ram et al. 2006b for details). The parameter space Θ is assumed to be an open subset of $\mathbb{R}^n$ and the detector that is used to capture the photons is denoted as $\mathcal{C}$, where $\mathcal{C} \subseteq \mathbb{R}^2$ is open. The temporal part of $\mathcal{G}$ is modeled as an inhomogeneous Poisson process with intensity Λ_θ called the *photon detection rate* and the spatial part of $\mathcal{G}$ is modeled as a sequence of mutually independent random variables with densities $\{f_{\theta,\tau}\}_{\tau \geq t_0}$ called the *photon distribution profile*. It is assumed that the spatial and temporal components are mutually independent of each other and that $f_{\theta,\tau}$ satisfies the regularity conditions necessary for the calculation of the Fisher information matrix (Ram et al. 2006b; Kay 1993).

The general expression of the Fisher information matrix for the image detection process $\mathcal{G}$ is given by (Ram et al. 2006b)

$$\mathbf{I}(\theta) = \int_{t_0}^{t} \int_{\mathcal{C}} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T \frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} dr d\tau, \quad \theta \in \Theta, \quad (1)$$

where $[t_0, t]$ denotes the time interval during which the data is acquired and the integration variable r denotes the 2D Cartesian coordinates (x, y). In the above equation, no specific assumptions have been made regarding the functional form of $f_{\theta,\tau}$ or Λ_θ. Therefore, the above expression of $\mathbf{I}(\theta)$ is applicable to a wide variety of imaging conditions, such as coherent/incoherent/partially-coherent light sources, polarized illumination and detection, etc. We note that the above equation is applicable to both stationary and moving objects, since we allow the density $f_{\theta,\tau}$, which describes the image profile of the object, to vary in time.

In order to quantify and compare the performance of the various experimental approaches considered in this paper, we make use of the Cramer-Rao inequality (Rao 1965), which states that for any unbiased estimator $\hat{\theta}$ of a $n \times 1$ vector parameter θ, $\mathrm{Cov}(\hat{\theta}) \geq \mathbf{I}^{-1}(\theta), \theta \in \Theta$, where $\mathbf{I}(\theta)$ denotes the Fisher information matrix and it is assumed that the inverse exists. From this inequality, it immediately follows that the ith leading diagonal entry of the inverse Fisher information matrix ($[\mathbf{I}^{-1}(\theta)]_{ii}$) provides a lower bound to the variance of the estimates of the ith component of the parameter vector (θ_i), $i = 1, \ldots, n$.

Throughout the paper, we adopt a parameterization in which the location of the two point sources are specified in terms of their Cartesian coordinates, i.e., (x_{01}, y_{01}) and (x_{02}, y_{02}). Hence the expressions for the Fisher information matrix will be given in terms of this parameterization. As we will see in subsequent sections, this parameterization not only simplifies the derivation of the Fisher information matrix for the different experimental approaches considered here, but it also helps in the analysis of the relationship between the distance estimation problem and the localization accuracy problem. To derive the Cramer-Rao lower bound for the distance parameter d, we require the analytical expression for the (inverse) Fisher information matrix of d. For this, we make use of the following coordinate transformation formula (Kay 1993)

$$\mathbf{I}^{-1}(d) = \left(\frac{\partial d}{\partial \theta} \right) \mathbf{I}^{-1}(\theta) \left(\frac{\partial d}{\partial \theta} \right)^T, \quad d \in [0, \infty), \quad (2)$$

where $\theta = (x_{01}, y_{01}, x_{02}, y_{02})$, $\mathbf{I}^{-1}(\theta)$ denotes the inverse Fisher information matrix corresponding to θ, and

$$\left(\frac{\partial d}{\partial \theta} \right)^T := \frac{1}{d} \begin{pmatrix} -(x_{02} - x_{01}) \\ -(y_{02} - y_{01}) \\ (x_{02} - x_{01}) \\ (y_{02} - y_{01}) \end{pmatrix}, \quad \theta \in \Theta.$$

3 Fisher information matrix for the simultaneous detection approach

In the simultaneous detection approach, the acquired image is assumed to contain the signal from both objects. Hence the photon detection rate Λ_θ and the photon distribution profile $f_{\theta,\tau}$ can be written as

$$\Lambda_\theta(\tau) = \Lambda_{\theta,1}(\tau) + \Lambda_{\theta,2}(\tau), \quad \theta \in \Theta, \quad \tau \geq t_0, \quad (3)$$

$$f_{\theta,\tau}(r) = \epsilon_{\theta,1}(\tau) f_{\theta,\tau,1}(r) + \epsilon_{\theta,2}(\tau) f_{\theta,\tau,2}(r), \quad r = (x, y) \in \mathcal{C}, \quad \theta \in \Theta, \quad \tau \geq t_0, \tag{4}$$

where $\mathcal{C}$ denotes the detector, $\Lambda_{\theta,1}$, $\Lambda_{\theta,2}$ and $f_{\theta,\tau,1}$, $f_{\theta,\tau,2}$ denote the photon detection rates and the photon distribution profiles of the two objects, respectively, and $\epsilon_{\theta,i}(\tau) := \Lambda_{\theta,i}(\tau)/(\Lambda_{\theta,1}(\tau) + \Lambda_{\theta,2}(\tau))$, $\theta \in \Theta$, $\tau \geq t_0$, $i = 1, 2$.

The results in this section are divided into two parts. In Sect. 3.1, we first derive general expressions of the Fisher information matrix for the simultaneous detection approach (Theorem 1). Here, we make no assumptions regarding the specific functional form of the photon detection rates $\Lambda_{\theta,i}$ or the photon distribution profiles $f_{\theta,\tau,i}$, $i = 1, 2$. Hence these results provide a general framework that is applicable to a wide variety of imaging scenarios.

In Sect. 3.2, we consider a concrete scenario (spatially invariant case) in optical microscopy where we assume a specific functional form for the photon distribution profiles $f_{\theta,\tau,i}$, $i = 1, 2$, which are expressed as a scaled and shifted version of the image of the objects. We then derive the Fisher information matrix for this functional form of $f_{\theta,\tau,i}$, $i = 1, 2$ (Theorem 2). As will be shown, the resulting Fisher information matrix can be expressed as a product decomposition of the form DCD^T, where D is an orthogonal matrix and C is a positive semidefinite matrix. Under weak assumptions of spatial symmetry for the image of the objects (which are typically satisfied in most situations), the product decomposition greatly simplifies the calculation of the Fisher information matrix and also facilitates the derivation of an analytical expression for the inverse Fisher information matrix (Corollary 1).

3.1 General expression of the Fisher information matrix

In many imaging applications, the unknown parameter vector θ can be expressed as $\theta = (\theta_f, \theta_\Lambda)$, where θ_f denotes the *spatial component* and θ_Λ denotes the *temporal component*. The spatial component θ_f typically consists of parameters that specify the location of one or more objects and the temporal component θ_Λ consists of parameters that specify the photon detection rates of the objects.

In the following theorem, we express the Fisher information matrix as a 2×2 block matrix. The terms in the leading diagonal (i.e., $\mathbf{S}_{sim}$ and $\mathbf{T}_{sim}$) correspond to the Fisher information matrix of the spatial θ_f and temporal θ_Λ components while the terms in the off-diagonal (i.e., $\mathbf{R}_{sim}$ and $\mathbf{R}_{sim}^T$) correspond to the coupling between the spatial and temporal components. We derive expressions for three practical scenarios. In the first scenario, we derive a general expression for the Fisher information matrix. In the second scenario, we consider the case where the photon detection rates are related to one another by a known scalar function β, i.e., $\beta(\tau)\Lambda_{\theta,1}(\tau) = \Lambda_{\theta,2}(\tau)$ for $\tau \geq t_0$ and $\theta \in \Theta$, where $\beta(\tau) \geq 0$. In some applications, the photon detection rates of the objects are assumed to be the same, i.e., $\Lambda_{\theta,1}(\tau) = \Lambda_{\theta,2}(\tau)$, $\tau \geq t_0$. We note that this condition is a special case of the second scenario considered here with $\beta(\tau) = 1$, $\tau \geq t_0$. For this scenario we show that the Fisher information matrix becomes block diagonal, which implies that the spatial θ_f and temporal θ_Λ components become decoupled. We note that this decoupling simplifies the subsequent analysis of the Fisher information matrix. In the third scenario, we assume that the photon distribution profiles of the objects are equal, i.e., $f_{\theta,\tau,1}(r) = f_{\theta,\tau,2}(r)$ for $r \in \mathcal{C}$, $\theta \in \Theta$ and $\tau \geq t_0$. This scenario arises in many applications, where the image profiles of the objects are assumed to be identical. For this scenario also we show that the Fisher information matrix becomes block diagonal.

Theorem 1 *Let $\Theta \subseteq \mathbb{R}^n$. For $\theta := (\theta_f, \theta_\Lambda) \in \Theta$, let $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$ be an image detection process, where Λ_θ and $f_{\theta,\tau}$ are defined in* Eqs. 3 and 4*, respectively. Assume that for $\theta \in \Theta$, $\tau \geq t_0$ and $i = 1, 2$,*

A1 $(\partial f_{\theta,\tau,i}(r)/\partial\theta_\Lambda) = 0, r \in \mathcal{C}$,
A2 $(\partial \Lambda_{\theta,i}(\tau)/\partial\theta_f) = 0$.

1. *Then the Fisher information matrix of $\mathcal{G}$ corresponding to the acquisition time interval $[t_0, t]$ for the simultaneous detection approach is given by*

$$\mathbf{I}_{sim}(\theta) = \begin{bmatrix} \mathbf{S}_{sim}(\theta) & \mathbf{R}_{sim}(\theta) \\ \mathbf{R}_{sim}^T(\theta) & \mathbf{T}_{sim}(\theta) \end{bmatrix}, \quad \theta \in \Theta,$$

where for $\theta \in \Theta$,

$$\mathbf{S}_{sim}(\theta) := \int_{t_0}^{t}\int_{\mathcal{C}} \frac{\Lambda_\theta(\tau)}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial\theta_f}\right)^T \frac{\partial f_{\theta,\tau}(r)}{\partial\theta_f} dr d\tau, \tag{5}$$

$$\mathbf{R}_{sim}(\theta) := \int_{t_0}^{t}\int_{\mathcal{C}} \frac{\Lambda_\theta(\tau)}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial\theta_f}\right)^T \frac{\partial f_{\theta,\tau}(r)}{\partial\theta_\Lambda} dr d\tau, \tag{6}$$

$$\begin{aligned}\mathbf{T}_{sim}(\theta) := &\int_{t_0}^{t} \frac{1}{\Lambda_\theta(\tau)} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial\theta_\Lambda}\right)^T \frac{\partial \Lambda_\theta(\tau)}{\partial\theta_\Lambda} d\tau \\ &+ \int_{t_0}^{t}\int_{\mathcal{C}} \frac{\Lambda_\theta(\tau)}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial\theta_\Lambda}\right)^T \frac{\partial f_{\theta,\tau}(r)}{\partial\theta_\Lambda} dr d\tau.\end{aligned} \tag{7}$$

2. *For $\beta(\tau) \geq 0, \tau \geq t_0$, assume, in addition to* **A1** *and* **A2**, *that*

A3 $\beta(\tau)\Lambda_{\theta,1}(\tau) = \Lambda_{\theta,2}(\tau), \tau \geq t_0$ *and* $\theta \in \Theta$.

Then the Fisher information matrix of $\mathcal{G}$ corresponding to the acquisition time interval $[t_0, t]$ for the simultaneous detection approach is given by

$$\mathbf{I}_{sim}(\theta) = \begin{bmatrix} \tilde{\mathbf{S}}_{sim}(\theta) & 0 \\ 0 & \tilde{\mathbf{T}}_{sim}(\theta) \end{bmatrix}, \quad \theta \in \Theta,$$

where for $\theta \in \Theta$,

$$\begin{aligned}\tilde{\mathbf{S}}_{sim}(\theta) := &\int_{t_0}^{t}\int_{\mathcal{C}} \frac{\Lambda_{\theta,1}(\tau)}{f_{\theta,\tau,1}(r) + \beta(\tau) f_{\theta,\tau,2}(r)} \left(\frac{\partial[f_{\theta,\tau,1}(r) + \beta(\tau) f_{\theta,\tau,2}(r)]}{\partial\theta_f}\right)^T \\ &\times \frac{\partial[f_{\theta,\tau,1}(r) + \beta(\tau) f_{\theta,\tau,2}(r)]}{\partial\theta_f} dr d\tau,\end{aligned}$$

$$\tilde{\mathbf{T}}_{sim}(\theta) := \int_{t_0}^{t} \frac{1+\beta(\tau)}{\Lambda_{\theta,1}(\tau)} \left(\frac{\partial \Lambda_{\theta,1}(\tau)}{\partial\theta_\Lambda}\right)^T \frac{\partial \Lambda_{\theta,1}(\tau)}{\partial\theta_\Lambda} d\tau.$$

3. *For $\theta \in \Theta$ and $\tau \geq t_0$, assume, in addition to* **A1** *and* **A2**, *that*

A4 $f_{\theta,\tau,1}(r) = f_{\theta,\tau,2}(r)$ *for* $r \in \mathcal{C}$.

Then the Fisher information matrix of $\mathcal{G}$ corresponding to the acquisition time interval $[t_0, t]$ for the simultaneous detection approach is given by

$$\mathbf{I}_{sim}(\theta) = \begin{bmatrix} \bar{\mathbf{S}}_{sim}(\theta) & 0 \\ 0 & \bar{\mathbf{T}}_{sim}(\theta) \end{bmatrix}, \quad \theta \in \Theta,$$

where for $\theta \in \Theta$,

$$\bar{\mathbf{S}}_{sim}(\theta) := \int_{t_0}^{t} \Lambda_\theta(\tau) d\tau \int_{\mathcal{C}} \frac{1}{f_{\theta,\tau,1}(r)} \left(\frac{\partial f_{\theta,\tau,1}(r)}{\partial \theta_f} \right)^T \frac{\partial f_{\theta,\tau,1}(r)}{\partial \theta_f} dr,$$

$$\bar{\mathbf{T}}_{sim}(\theta) := \int_{t_0}^{t} \frac{1}{\Lambda_\theta(\tau)} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda} \right)^T \frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda} d\tau.$$

Proof See Sect. A.1 in Appendix for proof. □

In many applications it is important to know whether the Fisher information matrix $\mathbf{I}(\theta)$ is (block) diagonal. For instance, it is well known that under certain conditions the maximum likelihood estimator of a vector parameter θ is asymptotically Gaussian distributed with mean θ and covariance $\mathbf{I}^{-1}(\theta)$ (see Van des Bos 2007). From the above Theorem, we see that if the photon detection rates can be expressed as a scalar function of one another or if the photon distribution profiles are identical, then $\mathbf{I}(\theta)$ becomes block diagonal. This implies that the maximum likelihood estimates of the spatial (θ_f) and temporal (θ_Λ) components of the unknown vector parameter θ are asymptotically independent. Moreover, if an efficient estimator of θ exists (i.e., an estimator whose covariance matrix is equal to $\mathbf{I}^{-1}(\theta), \theta \in \Theta$), then the estimates of θ_f and θ_Λ are uncorrelated. Another implication of block diagonality is that the Cramer-Rao lower bound of the spatial component θ_f is independent of the number of unknown parameters in the temporal component θ_Λ, and vice versa.

Remark 1 In result 2 of Theorem 1, we showed that the Fisher information matrix $\mathbf{I}_{sim}(\theta)$ is block diagonal if $\beta(\tau)\Lambda_{\theta,1}(\tau) = \Lambda_{\theta,2}(\tau)$ for $\tau \geq t_0$ and $\theta \in \Theta$, where $\beta(\tau) \geq 0$, $\tau \geq t_0$, is a known scalar function. We note that $\mathbf{I}_{sim}(\theta)$ will be block diagonal when $\Lambda_{\theta,1}(\tau) = \beta(\tau)\Lambda_{\theta,2}(\tau)$, $\tau \geq t_0$ and $\theta \in \Theta$ for $\beta(\tau) \geq 0$, $\tau \geq t_0$.

3.2 Fisher information matrix for the spatially invariant case

We next investigate a concrete scenario in optical microscopy where the image of the objects is spatially invariant, and we derive the Fisher information matrix for the simultaneous detection approach. Here, we introduce a specific parameterization of the spatial component θ_f of the parameter vector θ given by $\theta_f = \theta_c = (x_{01}, y_{01}, x_{02}, y_{02}) \in \Theta_c$, where (x_{01}, y_{01}) and (x_{02}, y_{02}) denote the Cartesian coordinates of the two objects, and Θ_c is the parameter space that is an open subset of $\mathbb{R}^4$. We consider the infinitely large detector $\mathcal{C} = \mathbb{R}^2$. For any given imaging condition, this infinite detector provides the best case scenario, where all the photons that reach the detector plane are detected.

In many microscopy applications, the image of an object can be considered to be invariant with respect to shifts in the object location (Young 1996). In the present context, the photon distribution profile $f_{\theta_c,\tau,i}$, $i = 1, 2$, can be expressed as a scaled and shifted version of the image of the object and is given by

$$f_{\theta_c,\tau,i}(r) = \frac{1}{M^2} q_i \left(\frac{x}{M} - x_{0i}, \frac{y}{M} - y_{0i} \right), \quad r = (x, y) \in \mathbb{R}^2, \tag{8}$$

where $\theta_c \in \Theta_c$, $\tau \geq t_0$, $i = 1, 2$, M denotes the total lateral magnification of the optical system, and q_i denotes the image function of the ith object, $i = 1, 2$. An image function q is

defined as the image of an object at unit magnification when the object is located at the origin of the coordinate axes. By definition, $f_{\theta_c,\tau,i}, i = 1, 2$, is assumed to satisfy the regularity conditions that are necessary for the calculation of the Fisher information matrix. Hence we impose appropriate conditions on the image functions, which are given in Definition 6 (see Appendix).

In many imaging experiments, the temporal component θ_Λ of the vector parameter θ is either assumed to be known or the photon detection rates are unknown but assumed to be equal ($\Lambda_{\theta,1}(\tau) = \Lambda_{\theta,2}(\tau), \tau \geq t_0$). In the former case, the Fisher information matrix of the simultaneous detection approach $\mathbf{I}_{sim}(\theta)$ trivially reduces to that of the spatial component θ_f i.e., $\mathbf{I}_{sim}(\theta) = \mathbf{S}_{sim}(\theta), \theta \in \Theta$. In the latter case, the Fisher information matrices of the spatial and temporal components are decoupled as shown in Result 2 of Theorem 1. Therefore in this section, we focus our analysis on the Fisher information matrix for the spatial component θ_f.

Without loss of generality, we assume that the photon detection rates of the objects are known, and hence we have

$$\Lambda_{\theta_c}(\tau) = \Lambda_1(\tau) + \Lambda_2(\tau), \quad \tau \geq t_0, \quad \theta_c \in \Theta_c, \tag{9}$$

where Λ_1 and Λ_2 denote the photon detection rates of the two objects. Further, the photon distribution profile $f_{\theta,\tau}$ is given by

$$f_{\theta_c,\tau}(r) := \epsilon_1(\tau) f_{\theta_c,\tau,1}(r) + \epsilon_2(\tau) f_{\theta_c,\tau,2}(r), \quad r \in \mathbb{R}^2, \quad \theta_c \in \Theta_c, \quad \tau \geq t_0, \tag{10}$$

where $\epsilon_i(\tau) = \Lambda_i(\tau)/(\Lambda_1(\tau) + \Lambda_2(\tau))$, and $f_{\theta_c,\tau,i}$ is given by Eq. 8 for $i = 1, 2, \tau \geq t_0$ and $\theta_c \in \Theta_c$.

In the next Theorem we derive an analytical expression of the Fisher information matrix for the spatial component θ_c pertaining to the specific parameterization of the photon detection rate Λ_{θ_c} and the photon distribution profile $f_{\theta_c,\tau}$ given in Eqs. 9 and 10, respectively. Here, we express the Fisher information matrix $\mathbf{S}_{sim}(\theta_c)$ as a 2×2 block matrix. As we shall see in Sect. 4, this expression will be used to analyze its relationship with the Fisher information matrix for the localization accuracy problem. We also derive a product decomposition for $\mathbf{S}_{sim}(\theta_c)$. This decomposition simplifies the calculation of the inverse of $\mathbf{S}_{sim}(\theta_c)$ and enables us to obtain an analytical expression for the same (Corollary 1).

Theorem 2 *Let $\Theta_c \subseteq \mathbb{R}^4$. For $\theta_c = (x_{01}, y_{01}, x_{02}, y_{02}) \in \Theta_c$, let $\mathcal{G}(\Lambda_{\theta_c}, \{f_{\theta_c,\tau}\}_{\tau \geq t_0}, \mathcal{C})$ be an image detection process, where Λ_θ and $f_{\theta,\tau}$ are given by* Eqs. 9 and 10*, respectively.*

1. *For $\theta_c \in \Theta_c$, the Fisher information matrix of the spatial component corresponding to the acquisition time interval $[t_0, t]$ for the simultaneous detection approach is given by*

$$\mathbf{S}_{sim}(\theta_c) = \begin{pmatrix} \mathbf{K}_{11}(\theta_c) & \mathbf{K}_{12}(\theta_c) \\ \mathbf{K}_{12}^T(\theta_c) & \mathbf{K}_{22}(\theta_c) \end{pmatrix}, \tag{11}$$

where for $\theta_c \in \Theta_c$ and $i, j = 1, 2$,

$$\mathbf{K}_{ij}(\theta_c) := \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{\Lambda_i(\tau)\Lambda_j(\tau)}{\Lambda_1(\tau) q_1(x - x_{01}, y - y_{01}) + \Lambda_2(\tau) q_2(x - x_{02}, y - y_{02})} \times \begin{pmatrix} \frac{\partial q_i(x-x_{0i}, y-y_{0i})}{\partial x} \frac{\partial q_j(x-x_{0j}, y-y_{0j})}{\partial x} & \frac{\partial q_i(x-x_{0i}, y-y_{0i})}{\partial x} \frac{\partial q_j(x-x_{0j}, y-y_{0j})}{\partial y} \\ \frac{\partial q_i(x-x_{0i}, y-y_{0i})}{\partial y} \frac{\partial q_j(x-x_{0j}, y-y_{0j})}{\partial x} & \frac{\partial q_i(x-x_{0i}, y-y_{0i})}{\partial y} \frac{\partial q_j(x-x_{0j}, y-y_{0j})}{\partial y} \end{pmatrix} dx dy d\tau. \tag{12}$$

2. *Let* $d = \sqrt{(x_{02}-x_{01})^2+(y_{02}-y_{01})^2}$ *and define* $\Theta_c^0 = \{(x_{01}, y_{01}, x_{02}, y_{02}) \mid (x_{01}, y_{01}) = (x_{02}, y_{02})\}$. *Then for* $\theta_c \in \Theta_c \backslash \Theta_c^0$, *the Fisher information matrix* $\mathbf{S}_{sim}(\theta_c)$ *given in result 1 of this Theorem can be written as*

$$\mathbf{S}_{sim}(\theta_c) = \mathbf{D}(\theta_c)\mathbf{C}(\theta_c)\mathbf{D}^T(\theta_c),$$

where for $\theta_c \in \Theta_c \backslash \Theta_c^0$

$$\mathbf{D}(\theta_c) := \begin{pmatrix} \tilde{\mathbf{D}}(\theta_c) & 0 \\ 0 & \tilde{\mathbf{D}}(\theta_c) \end{pmatrix}, \quad \tilde{\mathbf{D}}(\theta_c) := \frac{1}{d}\begin{pmatrix} x_{02}-x_{01} & -(y_{02}-y_{01}) \\ y_{02}-y_{01} & x_{02}-x_{01} \end{pmatrix}, \tag{13}$$

$$\mathbf{C}(\theta_c) := \begin{pmatrix} \mathbf{C}_{11}(\theta_c) & \mathbf{C}_{12}(\theta_c) \\ \mathbf{C}_{12}^T(\theta_c) & \mathbf{C}_{22}(\theta_c) \end{pmatrix}, \tag{14}$$

$$\mathbf{C}_{ij}(\theta_c) := \int_{t_0}^{t}\int_{\mathbb{R}^2} \frac{\Lambda_i(\tau)\Lambda_j(\tau)}{\Lambda_1(\tau)q_1(x+\frac{d}{2}, y) + \Lambda_2(\tau)q_2(x-\frac{d}{2}, y)} \times \begin{pmatrix} q'_{i,x}(x,y)q'_{j,x}(x,y) & q'_{i,x}(x,y)q'_{j,y}(x,y) \\ q'_{i,x}(x,y)q'_{j,y}(x,y) & q'_{i,y}(x,y)q'_{j,y}(x,y) \end{pmatrix} dxdyd\tau, \quad i,j=1,2, \tag{15}$$

with

$$q'_{i,\zeta}(x,y) := \begin{cases} \frac{\partial q_1(x+\frac{d}{2},y)}{\partial \zeta}, & i=1,\ (x,y)\in\mathbb{R}^2, \\ \frac{\partial q_2(x-\frac{d}{2},y)}{\partial \zeta}, & i=2,\ (x,y)\in\mathbb{R}^2, \end{cases} \quad \zeta \in \{x, y\}. \tag{16}$$

3. *Assume that* q_1 *and* q_2 *are symmetric along the* y *axis with respect to* $y = 0$, *i.e.,* $q_i(x,y) = q_i(x,-y)$, $(x,y) \in \mathbb{R}^2$ *and* $i = 1, 2$. *Then for* $\theta_c \in \Theta_c \backslash \Theta_c^0$ *and* $i = 1, 2$, $\mathbf{C}_{ij}(\theta_c)$ *is given by*

$$\mathbf{C}_{ij}(\theta_c) := \int_{t_0}^{t}\int_{\mathbb{R}^2} \frac{\Lambda_i(\tau)\Lambda_j(\tau)}{\Lambda_1(\tau)q_1(x+\frac{d}{2}, y) + \Lambda_2(\tau)q_2(x-\frac{d}{2}, y)} \times \begin{pmatrix} q'_{i,x}(x,y)q'_{j,x}(x,y) & 0 \\ 0 & q'_{i,y}(x,y)q'_{j,y}(x,y) \end{pmatrix} dxdyd\tau. \tag{17}$$

Proof Substituting for $f_{\theta_c,\tau}$ and Λ_{θ_c} in the expression for $\mathbf{I}_{ff}(\theta)$ given by Eq. 5 (see result 1 of Theorem 1) and using Lemma 2, we obtain result 1. For proof of results 2 and 3, please see Sect. A.2 in Appendix. □

In result 1 of the above Theorem, we obtained a block matrix representation of the Fisher information matrix $\mathbf{S}_{sim}(\theta_c)$. The leading diagonal terms correspond to the individual contributions from the two objects and the off-diagonal terms correspond to the coupling between the two objects. As we will show in the next Section, the coupling plays an important role in the analysis of the relationship between the Fisher information matrix for the simultaneous detection approach and that for the localization accuracy problem.

The product decomposition $\mathbf{D}(\theta_c)\mathbf{C}(\theta_c)\mathbf{D}^T(\theta_c)$ of $\mathbf{S}_{sim}(\theta_c)$ that we obtained in result 2 of the above Theorem has an interesting structure. The matrix $\mathbf{C}(\theta_c)$ is a special case of $\mathbf{S}_{sim}(\theta_c)$ where the y coordinates of the two objects are assumed to be the same, i.e., $y_{02} = y_{01}$, and the x coordinates of the two objects are equidistant from the origin. Note that the matrix $\mathbf{D}(\theta_c)$ is orthogonal (*i.e.*, $\mathbf{D}^{-1}(\theta_c) = \mathbf{D}^T(\theta_c)$). It should be pointed out that the product decomposition holds only when $(x_{01}, y_{01}) \neq (x_{02}, y_{02})$, i.e., when the distance d is not equal to zero, since

at $(x_{01}, y_{01}) = (x_{02}, y_{02})$ the matrix $\mathbf{D}(\theta_c)$ is not defined. An implication of this product decomposition is that for a given $\theta_c^s = (x_{01}^s, y_{01}^s, x_{02}^s, y_{02}^s)$ such that $(x_{01}^s, y_{01}^s) \neq (x_{02}^s, y_{02}^s)$, the Fisher information matrix for θ_c^s can be obtained by first computing the Fisher information matrix for $(-\frac{d}{2}, 0, \frac{d}{2}, 0)$ and then pre- and post-multiplying it with $\mathbf{D}(\theta_c^s)$ and $\mathbf{D}^T(\theta_c^s)$, respectively, where d denotes the distance between the two objects. In many practical situations, the image of the objects is symmetric along the y (and the x) axis. As shown in result 3 of Theorem 2, when this condition is satisfied, several entries of the matrix $\mathbf{C}(\theta_c)$ become zero, which in turn simplifies the calculation of $\mathbf{C}(\theta_c)$.

Remark 2 Consider the scenario when the distance between the two objects is zero, i.e. $x_{01} = x_{02}$ and $y_{01} = y_{02}$. For this scenario, the Fisher information matrix $\mathbf{S}_{sim}(\theta_c)$ given in result 1 of Theorem 2 is singular, if the photon detection rates and the image functions of the two objects are identical, i.e., $\Lambda_1 = \Lambda_2$ and $q_1 = q_2$ (also see Sect. 4.1). However, for distinct photon detection rates and image functions, $\mathbf{S}_{sim}(\theta_c)$ will, in general, be invertible even when the distance between the objects is zero.

In the following Corollary, we make use of the product decomposition of the Fisher information matrix $\mathbf{S}_{sim}(\theta_c)$ and the orthogonality of $\mathbf{D}(\theta_c)$ to obtain an analytical expression for the inverse of $\mathbf{S}_{sim}(\theta_c)$ when the distance d between the objects is non-zero.

Corollary 1 *Define* $\Theta_c^0 = \{(x_{01}, y_{01}, x_{02}, y_{02}) \mid (x_{01}, y_{01}) = (x_{02}, y_{02})\}$. *For* $\theta_c \in \Theta_c \backslash \Theta_c^0$, *let* $\mathbf{S}_{sim}(\theta_c)$ *be given by result 2 of* Theorem 2, $\mathbf{D}(\theta_c)$ *be given by* Eq. 13 *and* $\mathbf{C}_{ij}(\theta_c)$, $i = 1, 2$, *be given by* Eq. 17. *Assume that* q_1 *and* q_2 *are symmetric along the* y *axis with respect to* $y = 0$, *i.e.,* $q_i(x, y) = q_i(x, -y)$, $(x, y) \in \mathbb{R}^2$, $i = 1, 2$. *Then for* $\theta_c \in \Theta_c \backslash \Theta_c^0$, *we have*

$$\mathbf{S}_{sim}^{-1}(\theta_c) = \mathbf{D}(\theta_c)\mathbf{H}(\theta_c)\mathbf{D}^T(\theta_c),$$

where for $\theta_c \in \Theta_c \backslash \Theta_c^0$,

$$\begin{aligned}
\mathbf{H}(\theta_c) &= \begin{pmatrix} \Gamma(\theta_c) & 0 \\ 0 & \Gamma(\theta_c) \end{pmatrix} \begin{pmatrix} \mathbf{C}_{22}(\theta_c) & -\mathbf{C}_{12}(\theta_c) \\ -\mathbf{C}_{12}(\theta_c) & \mathbf{C}_{11}(\theta) \end{pmatrix} \begin{pmatrix} \Gamma(\theta_c) & 0 \\ 0 & \Gamma(\theta_c) \end{pmatrix}, \\
\Gamma(\theta_c) &:= \begin{pmatrix} \frac{1}{\sqrt{\Sigma_{11}(\theta_c)}} & 0 \\ 0 & \frac{1}{\sqrt{\Sigma_{22}(\theta_c)}} \end{pmatrix},
\end{aligned} \tag{18}$$

with

$$\Sigma_{ii}(\theta_c) := [\mathbf{C}_{11}(\theta_c)]_{ii}[\mathbf{C}_{22}(\theta_c)]_{ii} - ([\mathbf{C}_{12}(\theta_c)]_{ii})^2, \quad i = 1, 2, \quad \theta_c \in \Theta_c \backslash \Theta_c^0. \tag{19}$$

Proof The expression for $\mathbf{S}_{sim}^{-1}(\theta_c)$ is obtained by making use of the product decomposition of $\mathbf{S}_{sim}(\theta_c)$ and using the expression for the inverse of a block matrix (Zhang 1999). □

4 Simultaneous detection approach and the localization accuracy problem

In many optical microscopy applications, one of the central questions concerns the accuracy with which the location of a microscopic object (e.g., single molecule, biological subcellular structure such as a vesicle) can be determined, since this has several implications on the nature and type of studies that can be carried out (see Wong et al. 2011; Ober et al. 2004b). The Fisher information matrix for the problem of estimating the location of the ith object from its image is given by (see Ram et al. 2006b; Ober et al. 2004b)

$$\mathbf{Q}_i := \int_{t_0}^{t} \Lambda_i(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_i(x,y)} \begin{pmatrix} \left(\frac{\partial q_i(x,y)}{\partial x}\right)^2 & \frac{\partial q_i(x,y)}{\partial x}\frac{\partial q_i(x,y)}{\partial y} \\ \frac{\partial q_i(x,y)}{\partial x}\frac{\partial q_i(x,y)}{\partial y} & \left(\frac{\partial q_i(x,y)}{\partial y}\right)^2 \end{pmatrix} dxdy, \tag{20}$$

where $i = 1, 2$ and q_i and Λ_i denote the image function and the photon detection rate of the ith object, respectively, for $i = 1, 2$. The above equation was derived using the same stochastic framework used in this paper and it is assumed that the image contains signal from only the ith object, $i = 1, 2$.

In the following theorem we show how the Fisher information matrix $\mathbf{S}_{sim}(\theta_c)$ for the spatially invariant case of the simultaneous detection approach (Theorem 2) is related to the Fisher information matrix for the localization accuracy problem. Specifically, we show that when the distance tends to infinity, the Fisher information matrix $\mathbf{S}_{sim}(\theta_c)$ becomes equivalent to that of two independent localization accuracy problems.

Theorem 3 *For $\theta_c = (x_{01}, y_{01}, x_{02}, y_{02}) \in \Theta_c$, let $\mathbf{S}_{sim}(\theta_c)$ be given by result 1 of* Theorem 2. *For $i = 1, 2$, let $\mathbf{Q}_i$ be given by* Eq. 20. *Let Λ_1 and Λ_2, and q_1 and q_2 denote the photon detection rates and the image functions of the two objects, respectively. Assume that for $i = 1, 2, \zeta \in \{x, y\}$ and $y \in \mathbb{R}$,*

A1 $\lim_{x\to\pm\infty} q_i(x, y) = 0$,
A2 $\lim_{x\to\pm\infty} \frac{\partial q_2(x,y)}{\partial \zeta} = 0$.

Then

$$\mathbf{S}_{sim}^{inf} := \lim_{x_{02}\to\infty} \mathbf{S}_{sim}(\theta_c) = \lim_{x_{02}\to\infty} \begin{pmatrix} \mathbf{K}_{11}(\theta_c) & \mathbf{K}_{12}(\theta_c) \\ \mathbf{K}_{12}^T(\theta_c) & \mathbf{K}_{22}(\theta_c) \end{pmatrix} = \begin{pmatrix} \mathbf{Q}_1 & 0 \\ 0 & \mathbf{Q}_2 \end{pmatrix},$$

where $\mathbf{K}_{ij}(\theta_c)$, $i, j = 1, 2$ is given by Eq. 12.

Proof See Sect. A.3 in Appendix for proof. □

We would like to point out that in deriving the above result we assumed x_{02} to go to infinity. In general, the above result will hold when any one of the coordinates i.e., x_{01}, y_{01} or y_{02} is assumed to go to infinity. From the above Theorem we see that as the distance of separation becomes sufficiently large, the leading diagonal terms ($\mathbf{K}_{11}(\theta_c)$ and $\mathbf{K}_{22}(\theta_c)$) of the Fisher information matrix $\mathbf{S}_{sim}(\theta_c)$ for the simultaneous detection approach reduce to that of the localization accuracy problem for the two point sources (i.e., $\mathbf{Q}_1$ and $\mathbf{Q}_2$), and the off-diagonal term $\mathbf{K}_{12}(\theta_c)$ goes to zero. Note that the off-diagonal term represents the coupling between the two point sources.

From a practical standpoint, the knowledge of the behavior of the off-diagonal term as a function of the distance would enable the experimenter to determine whether it is necessary to calculate the full Fisher information matrix for the simultaneous detection approach or to only calculate the Fisher information matrix for the localization accuracy problem. As we will see in the next section the latter is typically much easier to calculate, since a closed form analytical expression can be obtained.

4.1 Example 1

We next illustrate the results derived in the prior sections by considering a specific image function and calculate the Fisher information matrix for the simultaneous detection approach and for the localization accuracy problem. Here, we make use of the Cramer-Rao inequality to obtain a lower limit to the accuracy (i.e., standard deviation) of the estimates of the

parameters of interest (see below). We assume the photon detection rates to be constant and equal i.e., $\Lambda_1(\tau) = \Lambda_2(\tau) = \Lambda_0$, $\tau \geq t_0$. We also assume the image functions to be identical and be given by the Airy profile, which, according to optical diffraction theory describes the image of an in-focus point source that is illuminated by incoherent, unpolarized light (Born and Wolf 1999). The analytical expression for the image functions can be written as

$$q_1(x, y) = q_2(x, y) := \frac{J_1^2\left(\frac{2\pi n_a}{\lambda}\sqrt{x^2+y^2}\right)}{\pi(x^2+y^2)}, \quad (x, y) \in \mathbb{R}^2, \tag{21}$$

where J_1 denotes the first order Bessel function of the first kind, $n_a > 0$ denotes the numerical aperture of the objective lens used to image the point source and $\lambda > 0$ denotes wavelength of the detected photons.

By making use of the Cramer-Rao inequality, we define three different quantities, namely the 2D fundamental resolution measure (FREM) for the simultaneous detection approach, the limit to the accuracy of the location coordinates for the simultaneous detection approach, and the fundamental limit to the localization accuracy. Then in Corollary 2, we consider two limiting cases of the distance parameter d, i.e., $d \to 0$ and $d \to \infty$, and derive analytical expressions of the 2D FREM for the simultaneous detection approach. In Sect. 4.1.1, we numerically calculate the above quantities for different values of d and discuss their implications.

Definition 1 The *2D FREM for the simultaneous detection approach* is defined as $\delta_d^{sim} := \sqrt{\mathbf{I}_{sim}^{-1}(d)}$, $d \in [0, \infty)$, where $\mathbf{I}_{sim}^{-1}(d)$ is obtained by substituting $\mathbf{S}_{sim}^{-1}(\theta_c)$ (Corollary 1) in the transformation formula given by Eq. 2.

Definition 2 The *limit to the accuracy of the location coordinates x_{0i} and y_{0j} for the simultaneous detection approach* are defined as $\delta_{x_{0i}}^{sim} := \sqrt{[\mathbf{S}_{sim}^{-1}(\theta_c)]_{(2i-1)(2i-1)}}$ and $\delta_{y_{0i}}^{sim} := \sqrt{[\mathbf{S}_{sim}^{-1}(\theta_c)]_{(2j)(2j)}}$, respectively, where $i, j = 1, 2$ and $\mathbf{S}_{sim}^{-1}(\theta_c)$ denotes the inverse Fisher information matrix given by Corollary 1 for $\theta_c = (x_{01}, y_{01}, x_{02}, y_{02}) \in \Theta_c$.

Definition 3 The *fundamental limit to the localization accuracy of the x-coordinate* of the ith object is defined as $\delta_x^{loc,i} := \sqrt{[\mathbf{Q}_i^{-1}]_{11}}$, $i = 1, 2$, and for the y-coordinate it is defined as $\delta_y^{loc,i} := \sqrt{[\mathbf{Q}_i^{-1}]_{22}}$, $i = 1, 2$, where $\mathbf{Q}_i$ is given by Eq. 20, for i = 1, 2.

For the specific image functions and photon detection rates considered in this example, it can be shown that (see Ober et al. 2004b).

$$\delta^{loc} = \delta_x^{loc,i} = \delta_y^{loc,i} := \frac{\lambda}{2\pi n_a\sqrt{\Lambda_0(t-t_0)}}, \quad i = 1, 2. \tag{22}$$

Corollary 2 *For $d \in [0, \infty)$, let δ_d^{sim} denote the 2D FREM for the simultaneous detection approach. For $i = 1, 2$, let Λ_i and q_i denote the photon detection rate and the image function of the ith object, respectively.*

1. *Assume that $q_1(x, y) = q_2(x, y), (x, y) \in \mathbb{R}^2$ and $\Lambda_1(\tau) = \lambda_2(\tau), \tau \geq t_0$. Then $lim_{d\to 0}\delta_d^{sim} = \infty$.*
2. *For $i = 1, 2$, assume that q_i is radially symmetric, i.e., there exists a q_i such that $q_i(x, y) := q_i(\sqrt{x^2+y^2}), (x, y) \in \mathbb{R}^2$ and $i = 1, 2$. Then*

$$\lim_{d\to\infty} \delta_d^{sim} = \sqrt{\left(\delta_{rs,1}^{loc}\right)^2 + \left(\delta_{rs,2}^{loc}\right)^2},$$

where for $i = 1, 2$,

$$\delta_{rs,i}^{loc} := \frac{1}{\sqrt{\pi \kappa_i \int_{t_0}^{t} \Lambda_i(\tau) d\tau}} \quad with \quad \kappa_i := \int_0^{\infty} \frac{1}{q_i(r)} \left(\frac{\partial q_i(r)}{\partial r} \right)^2 r dr. \tag{23}$$

1. *Let* δ^{loc} *be given by* Eq. 22. *For* $i = 1, 2$, *let* q_i *be an Airy profile that is given by* Eq. 21 *and* $\Lambda_1(\tau) = \Lambda_2(\tau) = \Lambda_0, \tau \geq t_0$. *Then* $\lim_{d \to \infty} \delta_d^{sim} = \sqrt{2}\delta^{loc}$.

Proof 1. By definition $\delta_d^{sim} = \sqrt{\mathbf{I}_{sim}^{-1}(d)}$, where $\mathbf{I}_{sim}^{-1}(d)$ is obtained by substituting $\mathbf{S}_{sim}^{-1}(\theta_c)$ (Corollary 1) in the transformation formula in Eq. 2. When $d \to 0$ then $x_{01} \to x_{02}$ and $y_{02} \to y_{02}$, and from Remark 2 it immediately follows that $\mathbf{S}_{sim}(\theta_c)$ is singular, where $\theta_c = (x_{01}, y_{01}, x_{02}, y_{02}) \in \Theta_c$ and $\mathbf{S}_{sim}(\theta_c)$ is given by Eq. 5. From this the result follows.

2. Without loss of generality, we assume that $d \to \infty$ implies $x_{02} \to \infty$. For $\theta_c = (x_{01}, y_{01}, x_{02}, y_{02}) \in \Theta_c$, consider the term $\mathbf{S}_{sim}(\theta_c)$ which is given by Eq. 5. Using Theorem 3 and Lemma 3 (see Appendix), we have

$$\lim_{x_{02} \to \infty} \mathbf{S}_{sim}(\theta_c) = \begin{bmatrix} \mathbf{Q}_1 & 0 \\ 0 & \mathbf{Q}_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\left(\delta_{rs,1}^{loc}\right)^2} \mathbf{1}_{2\times 2} & 0 \\ 0 & \frac{1}{\left(\delta_{rs,2}^{loc}\right)^2} \mathbf{I}_{2\times 2} \end{bmatrix}, \tag{24}$$

where $\mathbf{Q}_i, i = 1, 2$, denotes the Fisher information matrix for the localization accuracy problem (Eq. 20) and $\mathbf{1}_{2\times 2}$ denotes the 2×2 identity matrix. Define $\Delta_x := x_{02} - x_{01}$ and $\Delta_y := y_{02} - y_{01}$. Consider the term

$$\lim_{x_{02} \to \infty} \left(\frac{\partial d}{\partial \theta_c} \right)^T = \lim_{x_{02} \to \infty} \frac{1}{d} \begin{pmatrix} -(x_{02} - x_{01}) \\ -(y_{02} - y_{01}) \\ (x_{02} - x_{01}) \\ (y_{02} - y_{01}) \end{pmatrix} = \lim_{x_{02} \to \infty} \frac{1}{\sqrt{\Delta_x^2 + \Delta_y^2}} \begin{pmatrix} -\Delta_x \\ -\Delta_y \\ \Delta_x \\ \Delta_y \end{pmatrix}$$

$$= \lim_{x_{02} \to \infty} \begin{pmatrix} -\frac{1}{\sqrt{1 + \frac{\Delta_y^2}{\Delta_x^2}}} \\ -\frac{1}{\sqrt{\frac{\Delta_x^2}{\Delta_y^2} + 1}} \\ \frac{1}{\sqrt{1 + \frac{\Delta_y^2}{\Delta_x^2}}} \\ \frac{1}{\sqrt{\frac{\Delta_x^2}{\Delta_y^2} + 1}} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}. \tag{25}$$

Using Eqs. 24 and 25 in Eq. 2 and taking the limit $x_{02} \to \infty$, we have

$$\lim_{x_{02} \to \infty} \mathbf{I}_{sim}^{-1}(d) = \begin{pmatrix} -1 & 0 & 1 & 0 \end{pmatrix} \begin{bmatrix} \frac{1}{\left(\delta_{rs,1}^{loc}\right)^2} \mathbf{1}_{2\times 2} & 0 \\ 0 & \frac{1}{\left(\delta_{rs,2}^{loc}\right)^2} \mathbf{I}_{2\times 2} \end{bmatrix} \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$= \left(\delta_{rs,1}^{loc}\right)^2 + \left(\delta_{rs,2}^{loc}\right)^2.$$

From this the result follows.

3. The Airy profile given in Eq. 21 is radially symmetric. Hence substituting for q_i and $\Lambda_i, i = 1, 2$, in Eq. 23, we have $\delta^{loc}_{rs,1} = \delta^{loc}_{rs,2} = \delta^{loc}$ and from this the result immediately follows. □

4.1.1 Results

Here we numerically calculate the various quantities defined in Definitions 1–3. For this purpose, we assume the two point sources to be equidistant from the origin and to lie on a line segment that passes through the origin and subtends an angle of 45° with respect to the x-axis. We choose this specific configuration, since some of the calculated values (i.e., particular $\delta^{sim}_{x_{0i}}$ and $\delta^{sim}_{y_{0i}}, i = 1, 2$) become equal, which simplifies the presentation of the results.

Figure 2 shows the behavior of the 2D FREM δ^{sim}_d as a function of the distance of separation. The figure also shows the limit to the accuracy of x_{01} and x_{02} for the simultaneous detection approach, i.e., $\delta^{sim}_{x_{01}}$ and $\delta^{sim}_{x_{02}}$, respectively, (the result for y_{01} and y_{02} are analogous) as well as the fundamental limit to the localization accuracy δ^{loc} (Eq. 22). According to Rayleigh's resolution criterion, two identical point sources are said to be resolved in a microscope if their distance of separation is greater than or equal to $0.61\lambda/n_a$, where n_a denotes the numerical aperture of the microscope and λ denotes the wavelength of light emitted by the point sources. For the specific numerical values considered in Fig. 2, Rayleigh's resolution limit is ≈219 nm, and according to this criterion distances below 219 nm cannot resolved. In contrast, in Fig. 2 we see that the numerical value of the 2D FREM δ^{sim}_d is relatively small for a range of distances below the classical resolution limit of 219 nm. An immediate implication of this result is that if there exists an efficient estimator, then these distances can be determined with an accuracy as predicted by δ^{sim}_d.

Note that as the distance of separation becomes very small, δ^{sim}_d becomes numerically large thereby predicting poor accuracy in estimating the distance of separation. This is expected since under the assumptions of identical photon detection rates and image functions, when the distance d goes to zero the corresponding Fisher information matrix becomes singular and the 2D FREM δ^{sim}_d becomes infinitely large (result 1 of Corollary 2). As the distance of separation increases, δ^{sim}_d becomes smaller thereby predicting a relatively high accuracy in determining the distance between the two point sources. In particular, for large distances δ^{sim}_d approaches the fundamental limit to the localization accuracy δ^{loc}. This is expected, as it was shown in Theorem 3 that when $d \to \infty$, the Fisher information matrix for the simultaneous detection approach reduces to an expression that is equivalent to two independent localization accuracy problems. For the specific image functions considered here, $\delta^{sim}_d = \sqrt{2}\delta^{loc}$ in the limit $d \to \infty$ (result 2 of Corollary 2).

The results for $\delta^{sim}_{x_{01}}$ and $\delta^{sim}_{x_{02}}$ are also analogous to that of δ^{sim}_d. Note that although $\delta^{sim}_{x_{01}}$ ($\delta^{sim}_{x_{02}}$) and δ^{loc} provide lower bounds to the accuracy with which the x-coordinate of a point source can be determined, their behaviors are very different. In particular, $\delta^{sim}_{x_{01}}$ and $\delta^{sim}_{x_{02}}$ depend on the distance and become infinitely large in the limit $d \to 0$ (see Remark 2), whereas δ^{loc} is independent of the distance and remains finite for all values of d.

The above discussion raises the question that under what conditions $\delta^{sim}_{x_{01}}$ and $\delta^{sim}_{x_{02}}$, and more importantly δ^{sim}_d will remain finite as the distance goes to zero. In the next Section, we investigate this problem by considering a specical case of the simultaneous detection approach where we assume that one of the location coordinates is known. As we will see in Sect. 5.1, for this special case the limit to the accuracy of the distance d remains finite as $d \to 0$ for the specific image profiles and photon detection rates considered in the present example.

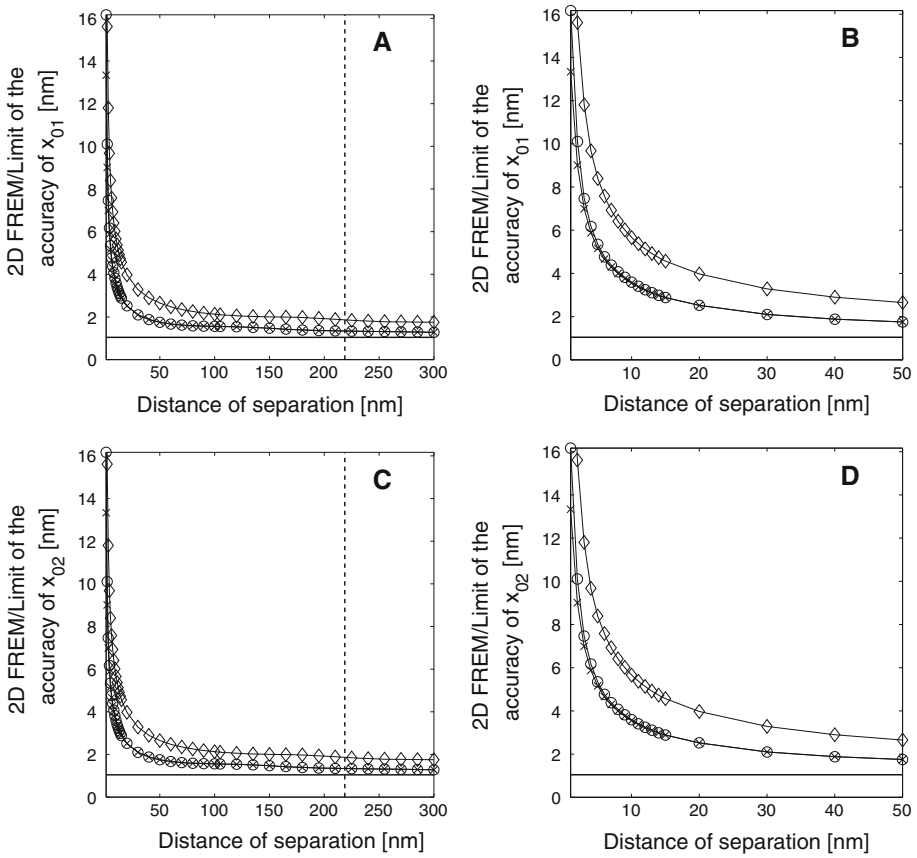

Fig. 2 Behavior of the 2D FREM δ_d^{sim} and the limit to the accuracy of x_{01} and x_{02}, i.e., $\delta_{x_{01}}^{sim}$ and $\delta_{x_{02}}^{sim}$, respectively, for the simultaneous detection approach. **a** shows δ_d^{sim} (*open diamond*) and $\delta_{x_{01}}^{sim}$ (*open circle*) (the results for $\delta_{y_{01}}^{sim}$ are similar) for a distance range of 1–300 nm, while **b** shows the same for a distance range of 1–50 nm. **c** shows δ_d^{sim} (*open diamond*) and $\delta_{x_{02}}^{sim}$ (*open circle*) (the results for $\delta_{y_{02}}^{sim}$ are similar) for a distance range of 1–300 nm, while **d** shows the same for a distance range of 1–50 nm. In all the panels, *solid line* denotes the fundamental limit to the localization accuracy δ^{loc} (Eq. 22). In **a**, **c**, the *vertical dashed line* denotes the classical Rayleigh's resolution limit, which is given by $0.61\lambda/n_a$. For all the plots, the numerical aperture is set to $n_a = 1.45$, the wavelength of the detected photons is set to $\lambda = 520$ nm, the photon detection rate for both point sources is set to $\Lambda_0 = 3000$ photons/s and the acquisition time interval is set to [0,1] s. For each value of distance, the location coordinates are set to $(x_{01}, y_{01}) = -(0.5d\cos\phi, 0.5d\sin\phi)$ and $(x_{02}, y_{02}) = (0.5d\cos\phi, 0.5d\sin\phi)$, with $\phi = \pi/4$ for all values of d. For the above numerical values, the Rayleigh's resolution limit is ≈219 nm

5 Special case of the simultaneous detection approach - location of one of the objects is known

It has been shown experimentally that distances well below the classical resolution criteria (e.g., Rayleigh's resolution criterion) can be resolved in a regular optical microscope when the location coordinates of one of the point sources is known a priori (Ram et al. 2006a; Gordon et al. 2004; Qu et al. 2004). For example, in a concrete experimental setting such a scenario arises when one wishes to study the interaction between a stationary object and a slow moving object. In many cases, the location coordinates of the stationary object can

be determined a priori (for instance from an image that only contains the stationary object) and therefore can be assumed to be known. Thus an important question then arises as to how accurately the distance between the two objects can be determined when the location of one of the objects is known. Here we address this problem by deriving the Fisher information matrix for this specific scenario.

For the present discussion, we assume that the acquired data consists of a pair of images, where one of the images contains photons from only one of the objects (for example, the stationary one) and the other image contains photons from both objects. Here, we assume that the location coordinates (x_{01}, y_{01}) of object 1 is determined from the first image and the location coordinates (x_{02}, y_{02}) of object 2 is determined from the second image. In the following Theorem, we derive the expression for the Fisher information matrix for the problem of estimating the location coordinates of the objects from such a pair of images. We assume that the photon detection rate of the objects is known. Further, we also assume the spatially invariant case (analogous to Sect. 3.2), where the photon distribution profile of the ith object $f_{\theta,\tau,i}, i = 1, 2$, is expressed as a scaled and shifted version of the image of that object (see Eq. 8).

As we will show, the Fisher information matrix reduces to that of two independent localization accuracy problems. We also show that the Fisher information matrix is invertible for all values of the location coordinates of the two objects including when the location coordinates are the same (i.e., when the distance equals zero).

Theorem 4 *Let $\Theta_c \subseteq \mathbb{R}^4$ be open. For $\theta_c = (x_{01}, y_{01}, x_{02}, y_{02}) \in \Theta_c, \tau \geq t_0$ and i = 1,2, let $f_{\theta_c,\tau,i}$ and Λ_i denote the photon distribution profile and the photon detection rate of the ith object, respectively, where $f_{\theta,\tau,i}$ is given by* Eq. 8. *For $\theta_c \in \Theta$ and $\tau \geq t_0$, let $\Lambda(\tau) := \Lambda_1(\tau) + \Lambda_2(\tau)$, and $f_{\theta_c,\tau}$ be given by* Eq. 10. *For $\theta_c \in \Theta_c$, let $\mathcal{G}_1(\Lambda_1, \{f_{\theta_c,\tau,1}\}_{\tau \geq t_0}, \mathbb{R}^2)$ and $\mathcal{G}_2(\Lambda, \{f_{\theta_c,\tau}\}_{\tau \geq t_0}, \mathbb{R}^2)$ denote two independent image detection processes.*

1. *Then for the two independent image detection processes $\mathcal{G}_1$ and $\mathcal{G}_2$, the Fisher information matrix of the spatial component corresponding to the acquisition time interval $[t_0, t]$ for the special case of the simultaneous detection approach is given by*

$$\mathbf{S}_{sim,sp}(\theta_c) := \begin{bmatrix} \mathbf{Q}_1 & 0 \\ 0 & \mathbf{K}_{22}(\theta_c) \end{bmatrix}, \quad \theta_c \in \Theta_c, \tag{26}$$

   where $\mathbf{Q}_1$ is given by Eq. 20 *and $\mathbf{K}_{22}(\theta_c), \theta_c \in \Theta_c$, is given by* Eq. 12.
2. *For $\theta_c \in \Theta_c$, $\mathbf{S}_{sim,sp}(\theta_c)$ is invertible including when $(x_{01}, y_{01}) = (x_{02}, y_{02})$.*

Proof See Sect. A.4 in Appendix for proof. □

From result 1 of the above Theorem we see that the Fisher information matrix for the special case of the simultaneous detection approach is a block diagonal matrix. The first term $\mathbf{Q}_1$ (Eq. 20) in the leading diagonal pertains to the Fisher information matrix for the localization accuracy problem corresponding to the location coordinates (x_{01}, y_{01}) of object 1. The second term $\mathbf{K}_{22}(\theta_c)$ (Eq. 12) in the leading diagonal is a component of the Fisher information matrix for the spatially invariant case of the simultaneous detection approach in which both location coordinates are unknown and are determined from a single image (Theorem 2). Importantly, this component $\mathbf{K}_{22}(\theta_c)$ is equivalent to the Fisher information matrix of the localization accuracy problem for the location coordinates (x_{02}, y_{02}) of object 2 in the presence of an extraneous background signal given by $\Lambda_1 q_1$, where Λ_1 and q_1 denote the photon detection rate and the image function of object 1, respectively. In this context, we would like to note that the effect of an extraneous background term on the localization

accuracy problem has been extensively investigated before (Ram et al. 2006b; Ober et al. 2004b).

In result 2 of the above Theorem, we showed that the Fisher information matrix is, in general, invertible for all values of the location coordinates of the two objects including when $(x_{01}, y_{01}) = (x_{02}, y_{02})$, i.e., when the distance between the two objects is zero. This is in contrast to the result obtained in Sect. 3.2, where we saw that the Fisher information matrix for the simultaneous detection approach becomes singular and therefore non-invertible when the distance is zero (assuming identical image profiles and photon detection rates; see Remark 2).

This brings out a very important aspect of the analyses carried out here. Specifically, the a priori knowledge of the location coordinates of one of the objects reduces the Fisher information matrix of the distance estimation problem to that of two independent localization accuracy problems. More importantly, it also removes the singularity of the Fisher information matrix when the distance is zero. The above result also explains the prior experimental observations of measuring nanometer scale distances well below the classical resolution criteria in a regular optical microscope when a priori information regarding the location coordinates of one of the objects is known (Gordon et al. 2004; Qu et al. 2004). In the next section, we further illustrate this through a specific example where we show that the CRLB of the distance parameter remains finite when the distance goes to zero.

We note that in the derivation of the above theorem, the Fisher information matrix for the second image only depends on the location coordinates of object 2, since it is assumed that the location of object 1 is known. However, since the second image contains signal from both objects, it also provides information about the location of object 1. Hence this can be used to improve the location estimates of object 1. A detailed analysis of such a scenario has been previously carried out by us, where, analogous to Theorem 4, we derived the Fisher information matrix for a pair of images but considered the case where both location coordinates were estimated from the second image (Ram et al. 2006a).

Remark 3 The results derived in the above Theorem pertains to the Fisher information matrix for the spatial component θ_f $(=\theta_c)$ of the unknown parameter vector θ, and we have assumed the temporal component θ_Λ of θ (and in turn the the photon detection rates of the objects) to be known. The above results will hold even if the temporal component θ_Λ is unknown provided the photon detection rates of the objects are related to one another through a scalar function β, i.e. $\Lambda_{\theta,1}(\tau) = \beta(\tau)\Lambda_{\theta,2}(\tau), \theta \in \Theta$ and $\tau \geq t_0$. This is due to the fact that under this condition, the Fisher information matrix for the spatial θ_f and temporal components θ_Λ are decoupled (see result 2 of Theorem 1). It should be pointed out that the assumption $\Lambda_{\theta,1} = \beta\Lambda_{\theta,2}, \theta \in \Theta$ is satisfied in many practical situations since the photon detection rates of the objects are typically assumed to be the same (i.e., $\beta = 1$).

5.1 Example 2

We now illustrate the results derived in the previous section by considering a specific image profile. Analogous to Sect. 4.1, we assume the image functions q_1 and q_2 to be identical Airy profiles given by Eq. 21 and set the photon detection rates to be constant and equal, i.e., $\Lambda_1(\tau) = \Lambda_2(\tau) = \Lambda_0, \tau \geq t_0$. We also define the 2D FREM for the special case of the simultaneous detection approach, which we denote as $\delta_d^{sim,sp}$. In Corollary 3, we consider two limiting cases of the distance parameter d, i.e., $d \to 0$ and $d \to \infty$ and derive analytical expressions for $\delta_d^{sim,sp}$ for the specific image functions and photon detection rates considered here.

Definition 4 The *2D FREM for the special case of the simultaneous detection approach* is defined as $\delta_d^{sim,sp} := \sqrt{\mathbf{I}_{sim,sp}^{-1}(d)}$, $d \in [0,\infty)$, where $\mathbf{I}_{sim,sp}(d)$ is obtained by substituting $\mathbf{S}_{sim,sp}^{-1}(\theta_c)$ (result 2 of Theorem 4) in the transformation formula given by Eq. 2.

Corollary 3 *For $d \in [0,\infty)$, let $\delta_d^{sim,sp}$ denote the 2D FREM for the special case of the simultaneous detection approach. For $i = 1, 2$, let Λ_i and q_i denote the photon detection rate and the image function of the ith object, respectively. Assume that $\Lambda_1(\tau) = \Lambda_2(\tau)$, $\tau \geq t_0$, $q_1(x,y) = q_2(x,y)$, $(x,y) \in \mathbb{R}^2$, and that q_1 is radially symmetric, i.e., there exists a q_1 such that $q_1(x,y) = q_1(\sqrt{x^2+y^2})$ for $(x,y) \in \mathbb{R}^2$.*

Then

1. $\lim_{d\to 0} \delta_d^{sim,sp} = \sqrt{3}\delta_{rs,1}^{loc}$,
2. $\lim_{d\to\infty} \delta_d^{sim,sp} = \sqrt{2}\delta_{rs,1}^{loc}$,
   where $\delta_{rs,1}^{loc}$ is given by Eq. 23.
3. *Let q_1 be an Airy profile that is given by* Eq. 21 *and $\Lambda_1(\tau) = \Lambda_0$, $\tau \geq t_0$. Then*

$$\lim_{d\to 0} \delta_d^{sim,sp} = \sqrt{3}\delta^{loc}, \quad \lim_{d\to\infty} \delta_d^{sim,sp} = \sqrt{2}\delta^{loc},$$

where δ^{loc} is given by Eq. 22.

Proof 1. By definition, q_1 is radially symmetric and hence from Lemma 3 it follows that $\mathbf{Q}_1^{-1} = (\delta_{rs,1}^{loc})^2 \mathbf{1}_{2\times 2}$, where $\mathbf{Q}_1$ denotes the Fisher information matrix for the localization accuracy problem of object 1 (Eq. 20) and $\mathbf{1}_{2\times 2}$ denotes the 2×2 identity matrix. Using this and Eq. 2, we get

$$\begin{aligned}(\delta_d^{sim,sp})^2 = \mathbf{I}_{sim,sp}^{-1}(d) &:= \frac{\partial d}{\partial \theta_c} \mathbf{S}_{sim,sp}^{-1}(\theta_c) \left(\frac{\partial d}{\partial \theta_c}\right)^T \\ &= \frac{1}{d^2} \begin{pmatrix} -(x_{02}-x_{01}) \\ -(y_{02}-y_{01}) \\ (x_{02}-x_{01}) \\ (y_{02}-y_{01}) \end{pmatrix}^T \begin{pmatrix} \mathbf{Q}_1^{-1} & 0 \\ 0 & \mathbf{K}_{22}^{-1}(\theta_c) \end{pmatrix} \begin{pmatrix} -(x_{02}-x_{01}) \\ -(y_{02}-y_{01}) \\ (x_{02}-x_{01}) \\ (y_{02}-y_{01}) \end{pmatrix} \\ &= (\delta_{rs,1}^{loc})^2 + \frac{1}{d^2} \begin{pmatrix} x_{02}-x_{01} & y_{02}-y_{01} \end{pmatrix} \mathbf{K}_{22}^{-1}(\theta_c) \begin{pmatrix} x_{02}-x_{01} \\ y_{02}-y_{01} \end{pmatrix}, \quad d \in [0,\infty),\end{aligned} \tag{27}$$

where $\mathbf{S}_{sim,sp}(\theta_c)$ is given by Eq. 26 and $\mathbf{K}_{22}(\theta_c)$ is given by Eq. 12. Because the photon detection rates and the image functions of the objects are assumed to be identical, we have $\mathbf{Q}_1 = \mathbf{Q}_2$, where $\mathbf{Q}_i$, $i = 1, 2$, is given by Eq. 20. Using this, we have

$$\begin{aligned}&\lim_{x_{01}\to x_{02}, y_{01}\to y_{02}} \mathbf{K}_{22}(\theta_c) \\ &= \lim_{x_{01}\to x_{02}, y_{01}\to y_{02}} \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{\Lambda_2^2(\tau)}{\Lambda_1(\tau) q_1(x-x_{01}, y-y_{01}) + \Lambda_2(\tau) q_2(x-x_{02}, y-y_{02})} \\ &\quad \times \begin{pmatrix} \left(\frac{\partial q_2(x-x_{02},y-y_{02})}{\partial x}\right)^2 & \frac{\partial q_2(x-x_{02},y-y_{02})}{\partial x} \frac{\partial q_2(x-x_{02},y-y_{02})}{\partial y} \\ \frac{\partial q_2(x-x_{02},y-y_{02})}{\partial y} \frac{\partial q_2(x-x_{02},y-y_{02})}{\partial x} & \left(\frac{\partial q_2(x-x_{02},y-y_{02})}{\partial y}\right)^2 \end{pmatrix} dx\,dy\,d\tau \\ &= \frac{\Lambda_0(t-t_0)}{2} \int_{\mathbb{R}^2} \frac{1}{q_2(x-x_{02}, y-y_{02})}\end{aligned}$$

$$\times \begin{pmatrix} \left(\frac{\partial q_2(x-x_{02},y-y_{02})}{\partial x}\right)^2 & \frac{\partial q_2(x-x_{02},y-y_{02})}{\partial x}\frac{\partial q_2(x-x_{02},y-y_{02})}{\partial y} \\ \frac{\partial q_2(x-x_{02},y-y_{02})}{\partial y}\frac{\partial q_2(x-x_{02},y-y_{02})}{\partial x} & \left(\frac{\partial q_2(x-x_{02},y-y_{02})}{\partial y}\right)^2 \end{pmatrix} dxdyd\tau$$

$$= \frac{1}{2}\mathbf{Q}_2 = \frac{1}{2}\mathbf{Q}_1 = \frac{1}{2(\delta_{rs,1}^{loc})^2}\mathbf{1}_{2\times 2}, \tag{28}$$

where we have used the shift-invariant property of Lebesgue integrals in the penultimate step. Define $\Delta_x := x_{02} - x_{01}$ and $\Delta_y := y_{02} - y_{01}$. Consider the term

$$\lim_{x_{01}\to x_{02}, y_{01}\to y_{02}} \frac{1}{d}\begin{pmatrix} x_{02}-x_{01} \\ y_{02}-y_{01}\end{pmatrix} = \lim_{x_{01}\to x_{02}}\lim_{y_{01}\to y_{02}} \begin{pmatrix} \frac{1}{\sqrt{1+\frac{\Delta_y^2}{\Delta_x^2}}} \\ \frac{1}{\sqrt{\frac{\Delta_x^2}{\Delta_y^2}+1}}\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}. \tag{29}$$

Substituting Eqs. 28 and 29 in Eq. 27 and taking the limit $d \to 0$, we get

$$\lim_{d\to 0}\left(\delta_d^{sim,sp}\right)^2 = \lim_{d\to 0}\mathbf{I}_{sim,sp}^{-1}(d) = \lim_{x_{01}\to x_{02}, y_{01}\to y_{02}}\mathbf{I}_{sim,sp}^{-1}(d) = \left(\delta_{rs,1}^{loc}\right)^2$$
$$+ \lim_{x_{01}\to x_{02}, y_{01}\to y_{02}} \frac{1}{d^2}(x_{02}-x_{01}\ \ y_{02}-y_{01})\mathbf{K}_{22}^{-1}(\theta_c)\begin{pmatrix} x_{02}-x_{01} \\ y_{02}-y_{01}\end{pmatrix}$$
$$= \left(\delta_{rs,1}^{loc}\right)^2 + 2\left(\delta_{rs,1}^{loc}\right)^2 (1\ 0)\,\mathbf{1}_{2\times 2}\begin{pmatrix}1\\0\end{pmatrix} = 3\left(\delta_{rs,1}^{loc}\right)^2.$$

From this the result immediately follows.

2. Proof is analogous to that of result 2 of Corollary 2.
3. The Airy profile given in Eq. 21 is radially symmetric. Substituting for q_1 and Λ_1 in results 1 and 2 of this Corollary, we get the desired results. □

Figure 3 shows the 2D FREM $\delta_d^{sim,sp}$ as a function of the distance for the special case of the simultaneous detection approach when the location coordinates (x_{01}, y_{01}) of one of the objects is assumed to be known. The figure also shows the 2D FREM for the simultaneous detection approach δ_d^{sim} when both location coordinates are assumed to be unknown (Sect. 1), and as a reference the fundamental limit to the localization accuracy δ^{loc} (Eq. 22). From the figure we see that as the distance of separation decreases, the δ_d^{sim} becomes infinitely large as $d \to 0$. In contrast, $\delta_d^{sim,sp}$ first increases but then decreases and then remains finite even when $d = 0$. In particular, for the specific image functions and photon detection rates considered here, $\delta_d^{sim,sp} = \sqrt{3}\delta^{loc}$ when $d = 0$ (result 1 of Corollary 3). An immediate implication of this result is that if the location coordinates of one of the objects is known, then it is possible to determine very small (nanometer scale) distances with relatively very high accuracy in an optical microscope. As the distance of separation increases, the 2D FREM $\delta_d^{sim,sp}$ behaves analogous to δ_d^{sim}. In particular, $\delta_d^{sim,sp} = \sqrt{2}\delta^{loc}$ when the distance becomes infinitely large. This implies that for very large distances of separation, the limit to the accuracy of estimating the distance is independent of the distance and is a constant.

We would like to point out that the analyses carried out in this Section have implications in a broader context of dealing with a singular Fisher information matrix, which represents a significant complication in the analysis of parameter estimation problems (e.g., see Stoica and Marzetta 2001). In particular our results illustrate how a priori information can be used to eliminate the singularity of the Fisher information matrix. It is important to note that

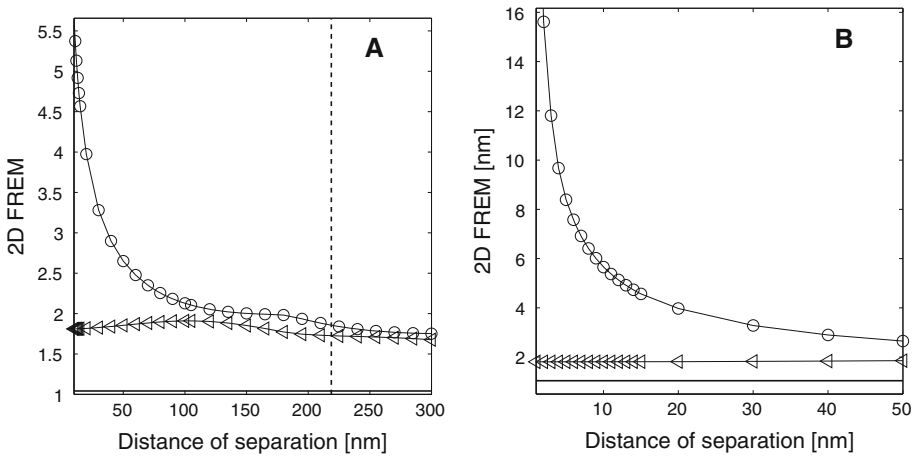

Fig. 3 Behavior of the 2D FREM $\delta_d^{sim,sp}$ for the special case of the simultaneous detection approach. **a** shows $\delta_d^{sim,sp}$ for a distance range of 10–300 nm for the special case of the simultaneous detection approach when the location coordinates (x_{01}, y_{01}) of object 1 is known (*pointing left triangle*). The panel also shows the 2D FREM δ_d^{sim} for the simultaneous detection approach when both location coordinates are unknown (*open circle*). **b** shows the same as **a** for a distance range of 1–50 nm. In all the panels, *solid line* denotes the fundamental limit to the localization accuracy δ^{loc} (Eq. 22), and in **a** the *vertical dashed line* denotes the Rayleigh's resolution limit. The numerical values used to generate the above plots are identical to those used in Fig. 2

the choice of a priori information intimately depends on the specifics of the experimental design, i.e., how the data is captured. This further underscores the importance of carrying out a rigorous analysis of the Fisher information matrix, as it provides the necessary insight into choosing the most appropriate experimental approach from the point of view of obtaining the best accuracy in estimating the parameters of interest.

6 Fisher information matrix for the separate detection approach

We next consider the case where the location coordinates of the two objects are independently estimated from two separate images. Such a scenario arises in a class of experimental techniques in which the photon emission from the objects are temporally separated (e.g., stochastic photoactivation (Betzig et al. 2006; Rust et al. 2006; Hess et al. 2006) and blinking (Lidke et al. 2005; Lagerholm et al. 2006)). In the following Theorem, we derive an analytical expression of the Fisher information matrix for the separate detection approach. Here, we assume the acquired data to consist of a pair of images, where the first image contains signal from only object 1 and the the second image contains signal from only object 2. As we will see, the Fisher information matrix for the separate detection approach will reduce to two independent localization accuracy problems.

Theorem 5 *Let* $\Theta_c \subseteq \mathbb{R}^4$ *be open. For* $\theta_c \in \Theta_c, \tau \geq t_0$ *and* $i = 1, 2$, *let* Λ_i *and* $f_{\theta_c,\tau,i}$ *denote the photon detection rate and the photon distribution profile of the* i^{th} *object, respectively, where* $f_{\theta_c,\tau,i}$ *is given by* Eq. 8. *For* $\theta_c \in \Theta_c$, *let* $\mathcal{G}_1(\Lambda_1, \{f_{\theta_c,\tau,1}\}_{\tau \geq t_0}, \mathbb{R}^2)$ *and* $\mathcal{G}_2(\Lambda_2, \{f_{\theta_c,\tau,2}\}_{\tau \geq t_0}, \mathbb{R}^2)$ *denote two independent image detection processes.*

1. *Then for the two independent image detection processes* $\mathcal{G}_1$ *and* $\mathcal{G}_2$, *the Fisher information matrix of the spatial component corresponding to the acquisition time interval* $[t_0, t]$ *for the separate detection approach is given by*

$$\mathbf{S}_{sep}(\theta_c) = \begin{bmatrix} \mathbf{Q}_1 & 0 \\ 0 & \mathbf{Q}_2 \end{bmatrix}, \quad \theta_c \in \Theta_c, \tag{30}$$

where $\mathbf{Q}_i$, $i = 1,2$, *is given in* Eq. 20

2. *For* $\theta_c \in \Theta_c$, $\mathbf{S}_{sep}(\theta_c)$ *is invertible including when* $(x_{01}, y_{01}) = (x_{02}, y_{02})$.

Proof Proof is analogous to that of Theorem 4. □

From the above Theorem, we see that the Fisher information matrix for the separate detection approach is block diagonal and is equivalent to two independent localization accuracy problems in the absence of any extraneous background signal. Note that the Fisher information matrix for the separate detection approach is independent of the location coordinates of the objects. This is in contrast to the simultaneous detection approach, where we saw that the Fisher information matrix depended on the location coordinates of the two objects (Theorem 2). In addition to this, for the simultaneous detection approach when both object coordinates are unknown the Fisher information matrix becomes block diagonal and reduces to that of two independent localization accuracy problems (in the absence of any extraneous background signal) only when the distance becomes infinitely large i.e., $d \to \infty$ (Theorem 3).

6.1 Example 3

To illustrate the result derived in this section, we consider a specific image function. Analogous to Sects. 4.1 and 5.1, we assume the image functions q_1 and q_2 to be identical Airy profiles given by Eq. 21 and set the photon detection rates to be constant and equal, i.e., $\Lambda_1(\tau) = \Lambda_2(\tau) = \Lambda_0, \tau \geq t_0$. We also define the 2D FREM for the separate detection approach δ_d^{sep}. Then in Corollary 4, we derive an analytical expression for δ_d^{sep} for the specific image functions and photon detection rates considered here.

Definition 5 The *2D FREM for the separate detection approach* is defined as $\delta_d^{sep} := \sqrt{\mathbf{I}_{sep}^{-1}(d)}, d \in [0, \infty)$, where $\mathbf{I}_{sep}^{-1}(d)$ is obtained by substituting $\mathbf{S}_{sep}^{-1}(\theta_c)$ (result 2 of Theorem 5) in the transformation formula given by Eq. 2.

Corollary 4 *For* $d \in [0, \infty)$, *let* δ_d^{sep} *denote the 2D FREM for the separate detection approach. For* $i = 1, 2$, *let* Λ_i *and* q_i *denote the photon detection rate and the image function of the* i*th object, respectively.*

1. *For* $i = 1, 2$, *assume that* q_i *is radially symmetric, i.e., there exists a* q_i *such that* $q_i(x, y) := q_i(\sqrt{x^2 + y^2}), (x, y) \in \mathbb{R}^2$ *and* $i = 1, 2$. *Then for* $d \in [0, \infty)$, *we have*

$$\delta_d^{sep} = \sqrt{\left(\delta_{rs,1}^{loc}\right)^2 + \left(\delta_{rs,2}^{loc}\right)^2},$$

where for $i = 1, 2$, $\delta_{rs,i}^{loc}$ *is given by* Eq. 23.

2. *For* $i = 1, 2$, *let* q_i *be an Airy profile that is given by* Eq. 21 *and* $\Lambda_1(\tau) = \Lambda_2(\tau) = \Lambda_0, \tau \geq t_0$. *Then for* $d \in [0, \infty)$, $\delta_d^{sep} = \sqrt{2}\delta^{loc}$, *where* δ^{loc} *is given by* Eq. 22.

Proof 1. Using Eq. 2 and Lemma 3, we have

$$\begin{aligned}(\delta_d^{sep})^2 = \mathbf{I}_{sep}^{-1}(d) &:= \frac{\partial d}{\partial \theta_c} \mathbf{S}_{sep}^{-1}(\theta_c) \left(\frac{\partial d}{\partial \theta_c}\right)^T \\ &= \frac{1}{d^2} \begin{pmatrix} -(x_{02} - x_{01}) \\ -(y_{02} - y_{01}) \\ (x_{02} - x_{01}) \\ (y_{02} - y_{01}) \end{pmatrix}^T \begin{pmatrix} \mathbf{Q}_1^{-1} & 0 \\ 0 & \mathbf{Q}_2^{-1} \end{pmatrix} \begin{pmatrix} -(x_{02} - x_{01}) \\ -(y_{02} - y_{01}) \\ (x_{02} - x_{01}) \\ (y_{02} - y_{01}) \end{pmatrix}\end{aligned}$$

$$= \frac{1}{d^2}\begin{pmatrix} -\Delta_x \\ -\Delta_y \\ \Delta_x \\ \Delta_y \end{pmatrix}^T \begin{pmatrix} \left(\delta_{rs,1}^{loc}\right)^2 & 0 & 0 & 0 \\ 0 & \left(\delta_{rs,1}^{loc}\right)^2 & 0 & 0 \\ 0 & 0 & \left(\delta_{rs,2}^{loc}\right)^2 & 0 \\ 0 & 0 & 0 & \left(\delta_{rs,2}^{loc}\right)^2 \end{pmatrix} \begin{pmatrix} -\Delta_x \\ -\Delta_y \\ \Delta_x \\ \Delta_y \end{pmatrix}$$

$$= \frac{1}{d^2}\left(\left(\Delta_x^2 + \Delta_y^2\right)\left(\delta_{rs,1}^{loc}\right)^2 + \left(\Delta_x^2 + \Delta_y^2\right)\left(\delta_{rs,2}^{loc}\right)^2\right)$$

$$= \left(\delta_{rs,1}^{loc}\right)^2 + \left(\delta_{rs,2}^{loc}\right)^2, \quad d \in [0, \infty),$$

where $\Delta_x = x_{02} - x_{01}$ and $\Delta_y = y_{02} - y_{01}$. From this the result immediately follows.

2. The Airy profile given in Eq. 21 is radially symmetric. Hence substituting for Λ_i and q_i, $i = 1, 2$, in result 1 of this Corollary, the result immediately follows. □

From the above result we see that the 2D FREM for the separate detection approach δ_d^{sep} is a constant and is independent of the distance of separation, if the image functions of the objects are radially symmetric. More specifically, when the image functions are assumed to be Airy profiles (Eq. 21), then the 2D FREM δ_d^{sep} is $\sqrt{2}$ times the fundamental limit to the localization accuracy δ^{loc}. This is in contrast to the simultaneous detection approach where the 2D FREM δ_d^{sim} (as well as $\delta_d^{sim,sp}$) depends on the distance and only in the limiting case when d becomes infinitely large, $\delta_d^{sim} = \sqrt{2}\delta^{loc}$ (Corollary 2). An immediate implication of the above result is that, if there exists an efficient estimator of the distance for the separate detection approach, then all distances can be determined with the same level of accuracy when the image profiles are radially symmetric.

7 Simulations

In the previous sections we investigated the Fisher information matrix of the distance d and calculated the 2D FREM for different experimental approaches. An important question then arises as to whether for a given experimental approach there exists an unbiased estimator that can attain the corresponding 2D FREM. In this section we address this question, where we use the Maximum Likelihood (ML) estimator to determine the distance d from simulated data and compare its performance (i.e. standard deviation) to the 2D FREM for the different experimental approaches. We consider all three approaches, i.e., the simultaneous detection approach, the special case of the simultaneous detection approach when one of the object locations are known, and the separate detection approach. We generate the acquired data through Monte-Carlo simulations which are discussed below. Here, we consider the data generation process for an ideal (non-pixelated) detector, where the acquired data consists of the spatial coordinates of the detected photons. We then use the maximum likelihood estimation algorithm on the simulated data to estimate the location coordinates of the objects, and from this we deduce the distance. Table 1 lists the standard deviations of the distance estimates for the different experimental approaches considered here. As we will see the ML estimator is unbiased and attains the 2D FREM for a range of distances when the sample size is sufficiently large.

Table 1 Results of the maximum likelihood estimator of the distance for the different experimental approaches considered here

Data set #	True value of distance nm	Mean distance estimates nm	SD of distance estimates nm	Resolution measure nm
A. Simultaneous detection approach				
1	10	10.22	5.87	5.89
2	20	20.01	4.14	4.23
3	50	49.99	2.67	2.65
4	100	99.99	2.14	2.12
5	200	200.06	1.97	1.93
6	500	500	1.64	1.68
B. Special case of the simultaneous detection approach				
1	10	10.15	1.56	1.81
2	20	20.03	1.68	1.82
3	50	50.03	1.85	1.85
4	100	100.02	1.91	1.91
5	200	200.03	1.78	1.74
6	500	500.01	1.56	1.6
C. Separate detection approach				
1	10	10.15	1.48	1.47
2	20	20.02	1.49	1.47
3	50	50.04	1.50	1.47
4	100	100.01	1.48	1.47
5	200	200.03	1.50	1.47
6	500	500.03	1.45	1.47

A shows the results for the simultaneous detection approach. B shows the results for the special case of the simultaneous detection approach, where one of the location coordinates is independently determined and is assumed to be known. C shows the results of the separate detection approach. The numerical values used to generate the data are identical to those used in Fig. 2. For all the data sets, the mean and standard deviation (SD) are obtained from 2,000 maximum likelihood estimates of the distance

7.1 Data simulation

We consider the two objects to be identical point sources. We set the photon detection rates of the two objects to be equal and constant, i.e. $\Lambda_{\theta,1}(\tau) := \Lambda_0, \tau \geq t_0$ and $\Lambda_{\theta,2}(\tau) := \Lambda_0, \tau \geq t_0$ and assume the image functions q_1 and q_2 to be identical Airy profiles given by Eq. 21. We generate a sequence of images $\{\mathcal{J}_{\theta,1}, \mathcal{J}_{\theta,2}, \ldots, \mathcal{J}_{\theta,N_{max}}\}$, where N_{max} denotes the total number of images. For $k = 1, \ldots, N_{max}$, the kth image is given by $\mathcal{J}_{\theta,k} := \{\mathcal{J}_{\theta,1,k}, \mathcal{J}_{\theta,2,k}\}$, where

$$\mathcal{J}_{\theta,i,k} := \{(x_1^{i,k}, y_1^{i,k}), (x_2^{i,k}, y_2^{i,k}), \ldots., (x_{N_{i,k}}^{i,k}, y_{N_{i,k}}^{i,k})\}, \quad i = 1, 2, \quad k = 1, \ldots, N_{max}, \quad (31)$$

denotes the signal from the ith object in the kth image for $k = 1, \ldots, N_{max}$ and $i = 1, 2$. In the above equation, $N_{i,k}$ denotes the number of detected photons from the ith object in the kth image for $i = 1, 2$ and $k = 1, \ldots, N_{max}$, and is a realization of the Poisson random variable with mean $\Lambda_0(t - t_0)$. The sequence $\{(x_m^{i,k}, y_m^{i,k}); m = 1, \ldots, N_{i,k})\}$ denotes the spatial coordinates of the detected photons from the ith object in the k^{th} image for

$i = 1, 2$ and $k = 1, \ldots, N_{max}$, and is a realization of $N_{i,k}$ random variables with density $f_{\theta_c,\tau,i}$ given by Eq. 8, which is generated by using a method described in Ober et al. (2004b).

7.2 Maximum likelihood estimator

For a general parameter estimation problem, the maximum likelihood estimator can be written as $\text{argmax}_\theta \ln(\mathcal{L}(\theta \mid \mathcal{Z})$ where $\mathcal{Z}$ denotes the data and $\mathcal{L}(\theta \mid \cdot)$ denotes the likelihood function. For the simultaneous detection approach, the acquired data pertaining to the kth image is given by $\mathcal{Z} = \mathcal{J}_{\theta,k} = \{\mathcal{J}_{\theta,1,k}, \mathcal{J}_{\theta,2,k}\}, k = 1, \ldots, N_{max}$ where $\mathcal{J}_{\theta,i,k}$ is defined in Eq. 31 and $\theta = \theta_c = (x_{01}, y_{01}, x_{02}, y_{02}) \in \mathbb{R}^4$.

For the special case of the simultaneous detection approach when one of the location coordinates are known, the acquired data consists of a pair of images $\{\mathcal{Z}_1, \mathcal{Z}_2\}$. We assume $\mathcal{Z}_1$ to be the image that contains the signal from object 1, i.e., $\mathcal{Z}_1 = \mathcal{J}_{\theta_1,1,k}$, and $\mathcal{Z}_2$ to be the image that contains the signal from both objects, i.e., $\mathcal{Z}_1 = \{\mathcal{J}_{\theta_1,1,k}, \mathcal{J}_{\theta_2,2,k}\}$, where $\mathcal{J}_{\theta_i,i,k}$ is defined in Eq. 31, $\theta_i = (x_{0i}, y_{0i}) \in \mathbb{R}^2, i = 1, 2$ and $k = 1, \ldots, N_{max}$. Here, we carry out two independent ML estimations on each image, i.e., $\text{argmax}_{\theta_1} \ln(\mathcal{L}(\theta_1 \mid \mathcal{Z}_1)$ and $\text{argmax}_{\theta_2} \ln(\mathcal{L}(\theta_2 \mid \mathcal{Z}_2, \hat{\theta}_1)$, where $\theta_i := (x_{0i}, y_{0i}) \in \mathbb{R}^2, i = 1, 2$. Note that while carrying out the maximum likelihood estimation with the second image $\mathcal{Z}_2$, we set the value of θ_1 to be equal to $\hat{\theta}_1$, where $\hat{\theta}_1$ denotes the maximum likelihood estimate of θ_1, which is determined from the first image.

For the separate detection approach, the acquired data consists of a pair of images $\{\mathcal{Z}_1, \mathcal{Z}_2\}$ each of which contains the image of only one of the objects. Here, we have $\mathcal{Z}_1 = \mathcal{J}_{\theta_1,1,k}$ and $\mathcal{Z}_2 = \mathcal{J}_{\theta_2,1,k}$, where $\mathcal{J}_{\theta_i,i,k}$ is defined in Eq. 31 for $\theta_i = (x_{0i}, y_{0i}) \in \mathbb{R}^2, i = 1, 2$ and $k = 1, \ldots, N_{max}$. For this approach, we carry out independent ML estimations on each image, i.e. $\text{argmax}_{\theta_1} \ln(\mathcal{L}(\theta_1 \mid \mathcal{Z}_1)$ and $\text{argmax}_{\theta_2} \ln(\mathcal{L}(\theta_2 \mid \mathcal{Z}_2)$, where $\theta_i := (x_{0i}, y_{0i}) \in \mathbb{R}^2, i = 1, 2$.

In all the three imaging scenarios, the ML estimates are determined computationally by using a gradient based optimization algorithm (`fminunc`) in the MATLAB programming language.

7.3 Comparison of ML estimator performance to the 2D FREM

Table 1 shows the results of the ML estimator for the different experimental approaches considered here. The table lists mean and standard deviation of the distance estimates as well as the 2D FREM of the distance for the corresponding experimental approach. From the table we see that for all the experimental approaches considered here, the mean value of the distance estimates is very close to the true value suggesting that the ML estimator is unbiased. Moreover, for a range of distances, the standard deviation of the distance is also consistently close to the 2D FREM thereby suggesting that the ML estimator is capable of achieving the theoretically best possible accuracy provided the sample size is sufficiently large. Note that the standard deviation of the ML estimates for the separate detection approach is almost a constant for a range of distances in agreement with the 2D FREM, which in turn shows that different distances can be estimated with the same level of accuracy.

A comparison of the standard deviations of the distance estimates (as well as the 2D FREMs) for the three approaches shows that for a range of distances considered in Table 1, the separate detection approach provides the best accuracy (i.e., the smallest 2D FREM/standard deviation) for determining the distance, followed by the special case of the simultaneous detection approach, and then followed by the simultaneous detection approach.

Acknowledgments This research was supported in part by the National Institutes of Health (R01 GM085575) and by a postdoctoral fellowship to S. R. from the National Multiple Sclerosis Society (FG-1798-A-1).

A Appendix

Definition 6 A function $q : \mathbb{R}^2 \to [0, \infty)$ is said to be an image function if the following properties are satisfied (see Ram et al. 2006b, p. 37).

1. $\int_{\mathbb{R}^2} q(x, y)dxdy = 1$,
2. $\frac{\partial q(x,y)}{\partial x}$ and $\frac{\partial q(x,y)}{\partial y}$ exist for every $(x, y) \in \mathbb{R}^2$,
3. $\int_{\mathbb{R}^2} \left|\frac{\partial q(x,y)}{\partial x}\right| dxdy < \infty$, $\int_{\mathbb{R}^2} \left|\frac{\partial q(x,y)}{\partial y}\right| dxdy < \infty$, and
4. $\int_{\mathbb{R}^2} \frac{1}{q(x,y)} \left(\frac{\partial q(x,y)}{\partial x}\right)^2 dxdy < \infty$, $\int_{\mathbb{R}^2} \frac{1}{q(x,y)} \left(\frac{\partial q(x,y)}{\partial y}\right)^2 dxdy < \infty$, and $\int_{\mathbb{R}^2} \frac{1}{q(x,y)} \frac{\partial q(x,y)}{\partial x} \frac{\partial q(x,y)}{\partial y} dxdy < \infty$.

Lemma 1 *For $\theta = (\theta_f, \theta_\Lambda) \in \Theta$, $\tau \geq t_0$ and $i = 1, 2$, let $f_{\theta,\tau,i}$ and $\Lambda_{\theta,i}$ denote the photon distribution profile and the photon detection rate of the ith object, respectively, and let Λ_θ and $f_{\theta,\tau}$ be given by* Eqs. 3 *and* 4*, respectively. Let $\mathcal{C}$ denote the detector.*

1. *For $\theta \in \Theta$ and $\tau \geq t_0$, if $f_{\theta,\tau,1}(r) = f_{\theta,\tau,2}(r), r \in \mathcal{C}$, then $\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} = 0, \theta \in \Theta$, $\tau \geq t_0, r \in \mathcal{C}$.*
2. *For $\theta \in \Theta$ and $\tau \geq t_0$, if $\beta(\tau)\Lambda_{\theta,1}(\tau) = \Lambda_{\theta,2}(\tau)$ for some $\beta(\tau) \geq 0$ that is independent of θ, then $\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} = 0, \theta \in \Theta, \tau \geq t_0, r \in \mathcal{C}$.*

Proof 1. For $\theta \in \Theta$, $\tau \geq t_0$ and $i = 1, 2$, let $\epsilon_{\theta,i}(\tau) = \Lambda_{\theta,i}(\tau)/\Lambda_\theta(\tau)$. Consider the term

$$\begin{aligned}\frac{\partial \epsilon_{\theta,1}(\tau)}{\partial \theta_\Lambda} + \frac{\partial \epsilon_{\theta,2}(\tau)}{\partial \theta_\Lambda} &= \frac{\Lambda_\theta(\tau)\frac{\partial \Lambda_{\theta,1}(\tau)}{\partial \theta_\Lambda} - \Lambda_{\theta,1}(\tau)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}}{\Lambda_\theta^2(\tau)} \\ &\quad + \frac{\Lambda_\theta(\tau)\frac{\partial \Lambda_{\theta,2}(\tau)}{\partial \theta_\Lambda} - \Lambda_{\theta,2}(\tau)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}}{\Lambda_\theta^2(\tau)} \\ &= \frac{\Lambda_\theta(\tau)\left(\frac{\partial \Lambda_{\theta,1}(\tau)}{\partial \theta_\Lambda} + \frac{\partial \Lambda_{\theta,2}(\tau)}{\partial \theta_\Lambda}\right) - (\Lambda_{\theta,1}(\tau) + \Lambda_{\theta,2}(\tau))\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}}{\Lambda_\theta^2(\tau)} \\ &= \frac{\Lambda_\theta(\tau)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda} - \Lambda_\theta(\tau)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}}{\Lambda_\theta^2(\tau)} = 0, \quad \theta \in \Theta, \quad \tau \geq t_0, \end{aligned} \tag{32}$$

where we have used the fact that $\Lambda_\theta(\tau) := \Lambda_{\theta,1}(\tau) + \Lambda_{\theta,2}(\tau)$, $\tau \geq t_0$ and $\theta \in \Theta$. Consider the term

$$\begin{aligned}\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} &= \frac{\partial \epsilon_{\theta,1}(\tau)}{\partial \theta_\Lambda} f_{\theta,\tau,1}(r) + \epsilon_{\theta,1}(\tau)\frac{\partial f_{\theta,\tau,1}(r)}{\partial \theta_\Lambda} + \frac{\partial \epsilon_{\theta,2}(\tau)}{\partial \theta_\Lambda} f_{\theta,\tau,2}(r) \\ &\quad + \epsilon_{\theta,2}(\tau)\frac{\partial f_{\theta,\tau,2}(r)}{\partial \theta_\Lambda}, \end{aligned} \tag{33}$$

where $\theta \in \Theta$, $\tau \geq t_0$ and $r \in \mathcal{C}$. Substituting for $f_{\theta,\tau,1}$ and $f_{\theta,\tau,2}$ in Eq. 33 and using Eq. 32, we have for $\theta \in \Theta$, $\tau \geq t_0$ and $r \in \mathcal{C}$

$$\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} = f_{\theta,\tau,1}(r)\left(\frac{\partial \epsilon_{\theta,1}(\tau)}{\partial \theta_\Lambda} + \frac{\partial \epsilon_{\theta,2}(\tau)}{\partial \theta_\Lambda}\right) = 0.$$

2. We have, $\epsilon_{\theta,1}(\tau) = \frac{1}{1+\beta(\tau)}$, $\theta \in \Theta$ and $\tau \geq t_0$, and $\epsilon_{\theta,2}(\tau) = \frac{\beta(\tau)}{1+\beta(\tau)}$, $\theta \in \Theta$ and $\tau \geq t_0$. Since $\beta(\tau)$ is independent of θ for $\tau \geq t_0$, $\frac{\partial \epsilon_{\theta,i}(\tau)}{\partial \theta_\Lambda} = 0, \theta \in \Theta, \tau \geq t_0$ and $i = 1, 2$. Substituting this in Eq. 33 the result follows. □

Lemma 2 *For $\theta_c = (x_{01}, y_{01}, x_{02}, y_{02}) \in \Theta_c$, $\tau \geq t_0$ and $i = 1, 2$, let $f_{\theta_c,\tau,i}$ be given by* Eq. 8. *Let $M > 0$. Then for $\theta_c \in \Theta_c$ and $\tau \geq t_0$, we have*

1. $\frac{\partial f_{\theta_c,\tau,i}(r)}{\partial x_{0i}} = -M \frac{\partial f_{\theta_c,\tau,i}(r)}{\partial x}$, $r = (x, y) \in \mathbb{R}^2$, $i = 1, 2$.
2. $\frac{\partial f_{\theta_c,\tau,i}(r)}{\partial y_{0i}} = -M \frac{\partial f_{\theta_c,\tau,i}(r)}{\partial y}$, $r = (x, y) \in \mathbb{R}^2$, $i = 1, 2$.

Proof 1. For $\theta_c = (x_{01}, x_{02}, y_{01}, y_{02}) \in \Theta_c$ and $i = 1, 2$, define $u_i := \frac{x}{M} - x_{0i}$ and $v_i := \frac{y}{M} - y_{0i}$. Then for $i = 1, 2$, we have

$$\begin{aligned}\frac{\partial f_{\theta_c,\tau,i}(r)}{\partial x_{0i}} &= \frac{1}{M^2}\frac{\partial q_i\left(\frac{x}{M} - x_{0i}, \frac{y}{M} - y_{0i}\right)}{\partial x_{0i}} = \frac{1}{M^2}\frac{\partial q_i(u_i, v_i)}{\partial u_i}\frac{\partial u_i}{\partial x_{0i}} = -\frac{1}{M^2}\frac{\partial q_i(u_i, v_i)}{\partial u_i} \\ &= \frac{1}{M^2}\frac{\partial q_i\left(\frac{x}{M} - x_{0i}, \frac{y}{M} - y_{0i}\right)}{\partial x}\frac{\partial x}{\partial u_i} = -M\frac{1}{M^2}\frac{\partial q_i\left(\frac{x}{M} - x_{0i}, \frac{y}{M} - y_{0i}\right)}{\partial x} \\ &= -M\frac{\partial f_{\theta_c,\tau,i}(r)}{\partial x},\end{aligned}$$

for $r = (x, y) \in \mathbb{R}^2$, $\theta_c \in \Theta_c$ and $\tau \geq t_0$.

2. Proof is similar to that of result 1. □

Lemma 3 *For $i = 1, 2$, let Q_i be given by* Eq. 20, *and Λ_i and q_i denote the photon detection rate and the image function of the ith object, respectively. For $i = 1, 2$, assume that q_i is radially symmetric with respect to the origin, i.e., there exists a q_i such that $q_i(x, y) = q_i(\sqrt{x^2 + y^2})$ for $(x, y) \in \mathbb{R}^2$ and $i = 1, 2$. Then for $i = 1, 2$,*

$$\mathbf{Q}_i = \frac{1}{(\delta^{loc}_{rs,i})^2}\mathbf{1}_{2\times 2},$$

where $\mathbf{1}_{2\times 2}$ denotes the 2×2 identity matrix and $\delta^{loc}_{rs,i}$, $i = 1, 2$, is given by Eq. 23.

Proof By definition, q_i, $i = 1, 2$, is symmetric along the x and y axes with respect to the origin. Using this, it can be shown that (see Ram et al. 2006b, p. 39)

$$\mathbf{Q}_i = \left(\int_{t_0}^{t} \Lambda_i(\tau)d\tau\right)\text{Diag}\left[\int_{\mathbb{R}^2}\frac{1}{q_i(x, y)}\left(\frac{\partial q_i(x, y)}{\partial x}\right)^2 dxdy \int_{\mathbb{R}^2}\frac{1}{q_i(x, y)}\left(\frac{\partial q_i(x, y)}{\partial y}\right)^2 dxdy\right],$$

where diag denotes the diagonal matrix. Further, using the fact that q_i, $i = 1, 2$, is radially symmetric, we have

$$[\mathbf{Q}_i]_{11} = \left(\int_{t_0}^{t} \Lambda_i(\tau)d\tau\right) \int_{\mathbb{R}^2} \frac{1}{q_i(x,y)} \left(\frac{\partial q_i(x,y)}{\partial x}\right)^2 dxdy$$

$$= \left(\int_{t_0}^{t} \Lambda_i(\tau)d\tau\right) \int_0^{2\pi}\int_0^{\infty} \frac{1}{q_i(r)} \left(\frac{\partial q_i(r)}{\partial r}\frac{\partial r}{\partial x}\right)^2 rdrd\phi$$

$$= \left(\int_{t_0}^{t} \Lambda_i(\tau)d\tau\right) \int_0^{2\pi} \cos^2(\phi)d\phi \int_0^{\infty} \frac{1}{q_i(r)} \left(\frac{\partial q_i(r)}{\partial r}\right)^2 rdr$$

$$= \left(\int_{t_0}^{t} \Lambda_i(\tau)d\tau\right) \left(\int_0^{2\pi} \frac{1+\cos(2\phi)}{2} d\phi\right) \kappa_i = \left(\int_{t_0}^{t} \Lambda_i(\tau)d\tau\right) \pi\kappa_i = \frac{1}{\left(\delta_{rs,i}^{loc}\right)^2},$$

where $i = 1, 2$, and κ_i is defined in Eq. 23. Similarly, we can show that for $i = 1, 2$, $[\mathbf{Q}_i]_{22} = 1/(\delta_{rs,i}^{loc})^2$. □

A.1 Proof of Theorem 1

Proof 1. Substituting for Λ_θ and $f_{\theta,\tau}$ in Eq. 1, and using assumptions **A1**–**A2** we get

$$\mathbf{I}_{sim}(\theta) = \int_{t_0}^{t}\int_{\mathcal{C}} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \begin{pmatrix} \Lambda_\theta(\tau)\left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_f}\right)^T \\ \Lambda_\theta(\tau)\left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda}\right)^T + f_{\theta,\tau}(r)\left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}\right)^T \end{pmatrix}$$

$$\times \begin{pmatrix} \Lambda_\theta(\tau)\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_f} & \Lambda_\theta(\tau)\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} + f_{\theta,\tau}(r)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda} \end{pmatrix} drd\tau$$

$$= \begin{bmatrix} \mathbf{S}_{sim}(\theta) & \int_{t_0}^{t}\int_{\mathcal{C}} \frac{1}{f_{\theta,\tau}(r)}\left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_f}\right)^T \left(\Lambda_\theta(\tau)\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} + f_{\theta,\tau}(r)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}\right) drd\tau \\ \left(\int_{t_0}^{t}\int_{\mathcal{C}} \frac{1}{f_{\theta,\tau}(r)}\left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_f}\right)^T \left(f_{\theta,\tau}(r)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda} + \Lambda_\theta(\tau)\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda}\right) drd\tau\right)^T & \int_{t_0}^{t}\int_{\mathcal{C}} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)}\left(\Lambda_\theta(\tau)\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} + f_{\theta,\tau}(r)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}\right)^T \left(\Lambda_\theta(\tau)\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} + f_{\theta,\tau}(r)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}\right) drd\tau \end{bmatrix}. \quad (34)$$

By definition, $f_{\theta,\tau}$ is a probability density function, which satisfies the regularity conditions that are necessary for the calculation of the Fisher information matrix (Kay 1993). Hence we have for $\theta \in \Theta$ and $\tau \geq t_0$,

$$\int_{\mathcal{C}} \frac{\partial f_{\theta,\tau}(r)}{\partial \theta} dr = \begin{pmatrix} \int_{\mathcal{C}} \frac{\partial f_{\theta,\tau}(r)}{\partial \theta_f} dr \\ \int_{\mathcal{C}} \frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} dr \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial \theta_f}\int_{\mathcal{C}} f_{\theta,\tau}(r)dr \\ \frac{\partial}{\partial \theta_\Lambda}\int_{\mathcal{C}} f_{\theta,\tau}(r)dr \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial \theta_f} 1 \\ \frac{\partial}{\partial \theta_\Lambda} 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (35)$$

Using Eq. 35, we have

$$
\begin{aligned}
[\mathbf{I}_{sim}(\theta)]_{12} &= [\mathbf{I}_{sim}(\theta)]_{21}^T \\
&= \int_{t_0}^{t}\int_{\mathcal{C}} \frac{1}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_f}\right)^T \left(\Lambda_\theta(\tau)\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} + f_{\theta,\tau}(r)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}\right) dr d\tau \\
&= \int_{t_0}^{t}\int_{\mathcal{C}} \frac{\Lambda_\theta(\tau)}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_f}\right)^T \frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} dr d\tau = \mathbf{R}_{sim}(\theta), \quad \theta \in \Theta. \qquad (36)
\end{aligned}
$$

Using Eq. 35 and the fact that $\int_{\mathcal{C}} f_{\theta,\tau}(r)dr = 1$ for $\theta \in \Theta$ and $\tau \geq t_0$, we have

$$
\begin{aligned}
[\mathbf{I}_{sim}(\theta)]_{22} &= \int_{t_0}^{t}\int_{\mathcal{C}} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\Lambda_\theta(\tau)\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} + f_{\theta,\tau}(r)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}\right)^T \\
&\quad \times \left(\Lambda_\theta(\tau)\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} + f_{\theta,\tau}(r)\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}\right) dr d\tau \\
&= \int_{t_0}^{t}\int_{\mathcal{C}} \frac{\Lambda_\theta(\tau)}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda}\right)^T \frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} dr d\tau \\
&\quad + \int_{t_0}^{t} \left(\int_{\mathcal{C}} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda}\right)^T dr\right) \frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda} d\tau \\
&\quad + \int_{t_0}^{t} \frac{1}{\Lambda_\theta(\tau)} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}\right)^T \frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda} d\tau \\
&\quad + \int_{t_0}^{t} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda}\right)^T \int_{\mathcal{C}} \frac{\partial f_{\theta,\tau}(r)}{\partial \theta_\Lambda} dr d\tau = \mathbf{T}_{sim}(\theta), \quad \theta \in \Theta. \qquad (37)
\end{aligned}
$$

Substituting Eqs. 36 and 37 in Eq. 34, the result immediately follows.

2. Using assumptions **A1** and **A3** it can be shown that $(\partial f_{\theta,\tau}(r)/\partial \theta_\Lambda) = 0, r \in \mathcal{C}, \theta \in \Theta, \tau \geq t_0$ (see result 3 of Lemma 1 in Appendix). Substituting this and using assumption **A3** in Eqs. 5, 6 and 7, we obtain the desired result.
3. Using assumptions **A1** and **A4** it can be shown that $(\partial f_{\theta,\tau}(r)/\partial \theta_\Lambda) = 0, r \in \mathcal{C}, \theta \in \Theta, \tau \geq t_0$ (see result 2 of Lemma 1 in Appendix). Further, by assumption **A4** we have $f_{\theta,\tau}(r) = f_{\theta,\tau,1}(r)(\epsilon_{\theta,1}(\tau) + \epsilon_{\theta,2}(\tau)) = f_{\theta,\tau,1}(r), r \in \mathcal{C}$ and $\tau \geq t_0$. Substituting these results in Eqs. 5, 6 and 7, we obtain the desired result. □

Proof of results 2 and 3 of Theorem 2

Proof 2. For $\theta_c \in \Theta_c \backslash \Theta_c^0$, define $s_x := (x_{01} + x_{02})/2, s_y := (y_{01} + y_{02})/2$ and $\phi = \tan^{-1}((y_{02} - y_{01})/(x_{02} - x_{01}))$. Then we have $x_{01} := s_x - \frac{d\cos\phi}{2}, y_{01} := s_y - \frac{d\sin\phi}{2}, x_{02} := s_x + \frac{d\cos\phi}{2}, y_{02} := s_y + \frac{d\sin\phi}{2}$. Substituting this in result 1 of the Theorem 2 and using the shift invariant property of Lebesgue intergrals, we get for $\theta_c \in \Theta_c \backslash \Theta_c^0$,

$$\mathbf{S}_{sim}(\theta_c) := \int_{t_0}^{t}\int_{\mathbb{R}^2} \frac{1}{\Lambda_1(\tau)q_1\left(x+\frac{d}{2}\cos\phi, y+\frac{d}{2}\sin\phi\right)+\Lambda_2(\tau)q_2\left(x-\frac{d}{2}\cos\phi, y-\frac{d}{2}\sin\phi\right)} \times \begin{bmatrix} \Lambda_1(\tau)\frac{\partial q_1\left(x+\frac{d}{2}\cos\phi, y+\frac{d}{2}\sin\phi\right)}{\partial x} \\ \Lambda_1(\tau)\frac{\partial q_1\left(x+\frac{d}{2}\cos\phi, y+\frac{d}{2}\sin\phi\right)}{\partial y} \\ \Lambda_2(\tau)\frac{\partial q_2\left(x-\frac{d}{2}\cos\phi, y-\frac{d}{2}\sin\phi\right)}{\partial x} \\ \Lambda_2(\tau)\frac{\partial q_2\left(x-\frac{d}{2}\cos\phi, y-\frac{d}{2}\sin\phi\right)}{\partial y} \end{bmatrix} \begin{bmatrix} \Lambda_1(\tau)\frac{\partial q_1\left(x+\frac{d}{2}\cos\phi, y+\frac{d}{2}\sin\phi\right)}{\partial x} \\ \Lambda_1(\tau)\frac{\partial q_1\left(x+\frac{d}{2}\cos\phi, y+\frac{d}{2}\sin\phi\right)}{\partial y} \\ \Lambda_2(\tau)\frac{\partial q_2\left(x-\frac{d}{2}\cos\phi, y-\frac{d}{2}\sin\phi\right)}{\partial x} \\ \Lambda_2(\tau)\frac{\partial q_2\left(x-\frac{d}{2}\cos\phi, y-\frac{d}{2}\sin\phi\right)}{\partial y} \end{bmatrix}^T dxdyd\tau. \tag{38}$$

For $(x, y) \in \mathbb{R}^2$, $\tau \geq t_0$ and $\theta_c \in \Theta_c \backslash \Theta_c^0$, let

$$Q_{\theta_c}^{+}(x, y, \tau) := \Lambda_1(\tau)q_1\left(x+\frac{d}{2}\cos\phi, y+\frac{d}{2}\sin\phi\right), \tag{39}$$

$$Q_{\theta_c}^{-}(x, y, \tau) := \Lambda_2(\tau)q_2\left(x-\frac{d}{2}\cos\phi, y-\frac{d}{2}\sin\phi\right). \tag{40}$$

For $\phi \in (0, 2\pi)$, define $T_\phi : \mathbb{R}^2 \to \mathbb{R}^2$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x\cos\phi + y\sin\phi \\ -x\sin\phi + y\cos\phi \end{pmatrix}.$$

The transformation T_ϕ maps the coordinates of a point on the 2D plane when the coordinate axes is rotated by an angle ϕ. Let $P^{\pm} := (x \pm \frac{d}{2}\cos\phi, y \pm \frac{d}{2}\sin\phi)$. Then

$$\tilde{P}^{\pm} := T_\phi P^{\pm} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \pm \frac{d}{2}\cos\phi \\ y \pm \frac{d}{2}\sin\phi \end{pmatrix} = \begin{pmatrix} x\cos\phi + y\sin\phi \pm \frac{d}{2} \\ -x\sin\phi + y\cos\phi \end{pmatrix}. \tag{41}$$

Using Eq. 41, we have for $\tau \geq t_0$ and $\theta_c \in \Theta_c \backslash \Theta_c^0$,

$$\begin{aligned} (Q_{\theta_c}^{+} \circ T_\phi)(x, y, \tau) &= \Lambda_1(\tau)q_1\left(T_\phi\left(x+\frac{d}{2}\cos\phi, y+\frac{d}{2}\sin\phi\right)\right) \\ &= \Lambda_1(\tau)q_1(T_\phi(P^+)) = \Lambda_1(\tau)q_1(\tilde{P}^+) \\ &= \Lambda_1(\tau)q_1\left(x\cos\phi + y\sin\phi + \frac{d}{2}, -x\sin\phi + y\cos\phi\right), \\ &\quad (x, y) \in \mathbb{R}^2, \end{aligned} \tag{42}$$

$$\begin{aligned} (Q_{\theta_c}^{-} \circ T_\phi)(x, y, \tau) &= \Lambda_2(\tau)q_2\left(x\cos\phi + y\sin\phi - \frac{d}{2}, -x\sin\phi + y\cos\phi\right), \\ &\quad (x, y) \in \mathbb{R}^2. \end{aligned} \tag{43}$$

Similarly, for $\theta_c \in \Theta_c \backslash \Theta_c^0$, $\tau \geq t_0$ and $\zeta \in \{x, y\}$,

$$\begin{aligned} \left(\frac{\partial Q_{\theta_c}^{+}}{\partial \zeta} \circ T_\phi\right)(x, y) &= \Lambda_1(\tau)\frac{\partial q_1(T_\phi(P^+))}{\partial \zeta} = \Lambda_1(\tau)\frac{\partial q_1(\tilde{P}^+)}{\partial \zeta} \\ &= \Lambda_1(\tau)\frac{\partial q_1\left(x\cos\phi + y\sin\phi + \frac{d}{2}, -x\sin\phi + y\cos\phi\right)}{\partial \zeta}, \\ &\quad (x, y) \in \mathbb{R}^2, \end{aligned} \tag{44}$$

$$\left(\frac{\partial Q^{-}_{\theta_c}}{\partial \zeta}\circ T_\phi\right)(x,y)=\Lambda_2(\tau)\frac{\partial q_2\left(x\cos\phi+y\sin\phi-\frac{d}{2},-x\sin\phi+y\cos\phi\right)}{\partial\zeta},$$
$$(x,y)\in\mathbb{R}^2. \tag{45}$$

By definition, the determinant of the Jacobian of T_ϕ is given by

$$\mathrm{Det}[T'_\phi]:=\mathrm{Det}\begin{bmatrix}\cos\phi & \sin\phi\\ sin\phi & \cos\phi\end{bmatrix}=1,\quad \phi\in(0,2\pi), \tag{46}$$

and for $(u,v):=T_\phi(x,y)$,

$$dudv=|\mathrm{Det}[T'_\phi]|dxdy=dxdy. \tag{47}$$

Substituting Eqs. 42–47 in the expression for $\mathbf{S}_{sim}(\theta_c)$ given in Eq. 38 and making use of the change of variables Theorem (Rudin 1987) we get,

$$\mathbf{S}_{sim}(\theta_c)=\int_{t_0}^{t}\int_{\mathbb{R}^2}\frac{1}{Q^{+}_{\theta_c}(x,y,\tau)+Q^{-}_{\theta_c}(x,y,\tau)}\begin{pmatrix}\frac{\partial Q^{+}_{\theta_c}(x,y,\tau)}{\partial x}\\ \frac{\partial Q^{+}_{\theta_c}(x,y,\tau)}{\partial y}\\ \frac{\partial Q^{-}_{\theta_c}(x,y,\tau)}{\partial x}\\ \frac{\partial Q^{-}_{\theta_c}(x,y,\tau)}{\partial y}\end{pmatrix}\begin{pmatrix}\frac{\partial Q^{+}_{\theta_c}(x,y,\tau)}{\partial x}\\ \frac{\partial Q^{+}_{\theta_c}(x,y,\tau)}{\partial y}\\ \frac{\partial Q^{-}_{\theta_c}(x,y,\tau)}{\partial x}\\ \frac{\partial Q^{-}_{\theta_c}(x,y,\tau)}{\partial y}\end{pmatrix}^{T}dxdyd\tau$$

$$=\int_{t_0}^{t}\int_{T_\phi(\mathbb{R}^2)}\left(\left(\frac{1}{Q^{+}_{\theta_c}+Q^{-}_{\theta_c}}\begin{pmatrix}\frac{\partial Q^{+}_{\theta_c}}{\partial x}\\ \frac{\partial Q^{+}_{\theta_c}}{\partial y}\\ \frac{\partial Q^{-}_{\theta_c}}{\partial x}\\ \frac{\partial Q^{-}_{\theta_c}}{\partial y}\end{pmatrix}\begin{pmatrix}\frac{\partial Q^{+}_{\theta_c}}{\partial x}\\ \frac{\partial Q^{+}_{\theta_c}}{\partial y}\\ \frac{\partial Q^{-}_{\theta_c}}{\partial x}\\ \frac{\partial Q^{-}_{\theta_c}}{\partial y}\end{pmatrix}^{T}\right)\circ T_\phi\right)(x,y,\tau)\mathrm{Det}|T'_\phi|dxdyd\tau$$

$$=\int_{t_0}^{t}\int_{\mathbb{R}^2}\frac{1}{Q^{+}_{\theta_c}(T_\phi(P^+))+Q^{-}_{\theta_c}(T_\phi(P^-))}\begin{pmatrix}\frac{\partial Q^{+}_{\theta_c}(T_\phi(P^+))}{\partial x}\\ \frac{\partial Q^{+}_{\theta_c}(T_\phi(P^+))}{\partial y}\\ \frac{\partial Q^{-}_{\theta_c}(T_\phi(P^-))}{\partial x}\\ \frac{\partial Q^{-}_{\theta_c}(T_\phi(P^-))}{\partial y}\end{pmatrix}\begin{pmatrix}\frac{\partial Q^{+}_{\theta_c}(T_\phi(P^+))}{\partial x}\\ \frac{\partial Q^{+}_{\theta_c}(T_\phi(P^+))}{\partial y}\\ \frac{\partial Q^{-}_{\theta_c}(T_\phi(P^-))}{\partial x}\\ \frac{\partial Q^{-}_{\theta_c}(T_\phi(P^-))}{\partial y}\end{pmatrix}^{T}dxdyd\tau$$

$$=\int_{t_0}^{t}\int_{\mathbb{R}^2}\frac{1}{Q^{+}_{\theta_c}(\tilde{P}^+)+Q^{-}_{\theta_c}(\tilde{P}^-)}\begin{pmatrix}\frac{\partial Q^{+}_{\theta_c}(\tilde{P}^+)}{\partial x}\\ \frac{\partial Q^{+}_{\theta_c}(\tilde{P}^+)}{\partial y}\\ \frac{\partial Q^{-}_{\theta_c}(\tilde{P}^-)}{\partial x}\\ \frac{\partial Q^{-}_{\theta_c}(\tilde{P}^-)}{\partial y}\end{pmatrix}\begin{pmatrix}\frac{\partial Q^{+}_{\theta_c}(\tilde{P}^+)}{\partial x}\\ \frac{\partial Q^{+}_{\theta_c}(\tilde{P}^+)}{\partial y}\\ \frac{\partial Q^{-}_{\theta_c}(\tilde{P}^-)}{\partial x}\\ \frac{\partial Q^{-}_{\theta_c}(\tilde{P}^-)}{\partial y}\end{pmatrix}^{T}dxdyd\tau$$

$$=\int_{t_0}^{t}\int_{\mathbb{R}^2}\frac{1}{\Lambda_1(\tau)q_1\left(x\cos\phi+y\sin\phi+\frac{d}{2},-x\sin\phi+y\cos\phi\right)+\Lambda_2(\tau)q_2\left(x\cos\phi+y\sin\phi-\frac{d}{2},-x\sin\phi+y\cos\phi\right)}$$
$$\times\begin{pmatrix}\Lambda_1(\tau)\frac{\partial q_1\left(x\cos\phi+y\sin\phi+\frac{d}{2},-x\sin\phi+y\cos\phi\right)}{\partial x}\\ \Lambda_1(\tau)\frac{\partial q_1\left(x\cos\phi+y\sin\phi+\frac{d}{2},-x\sin\phi+y\cos\phi\right)}{\partial y}\\ \Lambda_2(\tau)\frac{\partial q_2\left(x\cos\phi+y\sin\phi-\frac{d}{2},-x\sin\phi+y\cos\phi\right)}{\partial x}\\ \Lambda_2(\tau)\frac{\partial q_2\left(x\cos\phi+y\sin\phi-\frac{d}{2},-x\sin\phi+y\cos\phi\right)}{\partial y}\end{pmatrix}$$

$$\times \begin{pmatrix} \Lambda_1(\tau)\frac{\partial q_1\left(x\cos\phi+y\sin\phi+\frac{d}{2},-x\sin\phi+y\cos\phi\right)}{\partial x} \\ \Lambda_1(\tau)\frac{\partial q_1\left(x\cos\phi+y\sin\phi+\frac{d}{2},-x\sin\phi+y\cos\phi\right)}{\partial y} \\ \Lambda_2(\tau)\frac{\partial q_2\left(x\cos\phi+y\sin\phi-\frac{d}{2},-x\sin\phi+y\cos\phi\right)}{\partial x} \\ \Lambda_2(\tau)\frac{\partial q_2\left(x\cos\phi+y\sin\phi-\frac{d}{2},-x\sin\phi+y\cos\phi\right)}{\partial y} \end{pmatrix}^T dxdyd\tau$$

$$= \int_{t_0}^{t}\int_{\mathbb{R}^2} \frac{1}{\Lambda_1(\tau)q_1\left(u+\frac{d}{2},v\right)+\Lambda_2(\tau)q_2\left(u-\frac{d}{2},v\right)} \begin{pmatrix} \Lambda_1(\tau)\left(\cos\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial u}-\sin\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial v}\right) \\ \Lambda_1(\tau)\left(\sin\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial u}+\cos\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial v}\right) \\ \Lambda_2(\tau)\left(\cos\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial u}-\sin\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial v}\right) \\ \Lambda_2(\tau)\left(\sin\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial u}+\cos\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial v}\right) \end{pmatrix}$$

$$\times \begin{pmatrix} \Lambda_1(\tau)\left(\cos\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial u}-\sin\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial v}\right) \\ \Lambda_1(\tau)\left(\sin\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial u}+\cos\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial v}\right) \\ \Lambda_2(\tau)\left(\cos\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial u}-\sin\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial v}\right) \\ \Lambda_2(\tau)\left(\sin\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial u}+\cos\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial v}\right) \end{pmatrix}^T dudvd\tau, \quad \theta_c \in \Theta_c, \tag{48}$$

where $u := x\cos\phi + y\sin\phi$ and $v := -x\sin\phi + y\cos\phi$. Further, for $\theta_c \in \Theta_c \backslash \Theta_c^0$, $\tau \geq t_0$ and $(x, y) \in \mathbb{R}^2$, we have

$$\begin{pmatrix} \Lambda_1(\tau)\left(\cos\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial u}-\sin\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial v}\right) \\ \Lambda_1(\tau)\left(\sin\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial u}+\cos\phi\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial v}\right) \\ \Lambda_2(\tau)\left(\cos\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial u}-\sin\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial v}\right) \\ \Lambda_2(\tau)\left(\sin\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial u}+\cos\phi\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial v}\right) \end{pmatrix}$$

$$= \begin{bmatrix} \cos\phi & -\sin\phi & 0 & 0 \\ \sin\phi & \cos\phi & 0 & 0 \\ 0 & 0 & \cos\phi & -\sin\phi \\ 0 & 0 & \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \Lambda_1(\tau)\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial u} \\ \Lambda_1(\tau)\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial v} \\ \Lambda_2(\tau)\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial u} \\ \Lambda_2(\tau)\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial v} \end{bmatrix}$$

$$= \frac{1}{d}\begin{bmatrix} x_{02}-x_{01} & -(y_{02}-y_{01}) & 0 & 0 \\ y_{02}-y_{01} & x_{02}-x_{01} & 0 & 0 \\ 0 & 0 & x_{02}-x_{01} & -(y_{02}-y_{01}) \\ 0 & 0 & y_{02}-y_{01} & x_{02}-x_{01} \end{bmatrix} \begin{bmatrix} \Lambda_1(\tau)\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial u} \\ \Lambda_1(\tau)\frac{\partial q_1\left(u+\frac{d}{2},v\right)}{\partial v} \\ \Lambda_2(\tau)\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial u} \\ \Lambda_2(\tau)\frac{\partial q_2\left(u-\frac{d}{2},v\right)}{\partial v} \end{bmatrix}$$

$$= \mathbf{D}(\theta_c) \begin{pmatrix} \Lambda_1(\tau)q'_{1,x}(x,y) \\ \Lambda_1(\tau)q'_{1,y}(x,y) \\ \Lambda_2(\tau)q'_{2,x}(x,y) \\ \Lambda_2(\tau)q'_{2,y}(x,y) \end{pmatrix},$$

where $\mathbf{D}(\theta_c)$ is defined in Eq. 13, $q'_{i,\zeta}, i = 1, 2, \zeta \in \{x, y\}$ is given by Eq. 16 and we have used the fact that $\cos\phi := (x_{02} - x_{01})/d$ and $\sin\phi := (y_{02} - y_{01})/d$. Substituting the above expression in Eq. 48, the result immediately follows.

3. To prove this result we need to show that the off-diagonal terms of $\mathbf{C}_{ij}(\theta_c)$ are zero, for $i, j = 1, 2$ and $\theta_c \in \Theta_c \backslash \Theta_c^0$. For $\theta_c \in \Theta_c \backslash \Theta_c^0$, $\tau \geq t_0$ and $(x, y) \in \mathbb{R}^2$, let

$$W^1_{\theta_c}(x, y, \tau) := \Lambda_1(\tau) q_1\left(x + \frac{d}{2}, y\right), \quad W^2_{\theta_c}(x, y, \tau) := \Lambda_2(\tau) q_2\left(x - \frac{d}{2}, y\right). \tag{49}$$

Define $T_Y : \mathbb{R}^2 \times [t_0, \infty) \to \mathbb{R}^2 \times [t_0, \infty), (x, y, \tau) \mapsto (x, -y, \tau)$. Since q_1 and q_2 are symmetric along the y axis with respect to $y = 0$, we have $W^1_{\theta_c}(x, y, \tau) = (W^1_{\theta_c} \circ T_Y)(x, , y, \tau)$ and $W^2_{\theta_c}(x, y, \tau) = (W^2_{\theta_c} \circ T_Y)(x, , y, \tau)$ for $\theta_c \in \Theta_c \backslash \Theta_c^0$, $(x, y) \in \mathbb{R}^2$ and $\tau \geq t_0$. This implies that for $\theta_c \in \Theta_c \backslash \Theta_c^0$, $(x, y) \in \mathbb{R}^2$ and $\tau \geq t_0$, we have

$$\begin{aligned} U^{\pm}_{\theta_c}(x, y, \tau) &= \Lambda_1(\tau) q_1\left(x - \frac{d}{2}, y\right) \pm \Lambda_2(\tau) q_2\left(x + \frac{d}{2}, y\right) \\ &= \left(U^{\pm}_{\theta_c} \circ T_Y\right)(x, y, \tau), \end{aligned} \tag{50}$$

$$\frac{\partial W^i_{\theta_c}(x, y, \tau)}{\partial x} = \left(\frac{\partial W^i_{\theta_c}}{\partial x} \circ T_Y\right)(x, y, \tau), \quad i = 1, 2, \tag{51}$$

$$\frac{\partial W^i_{\theta_c}(x, y, \tau)}{\partial y} = -\left(\frac{\partial W^i_{\theta_c}}{\partial y} \circ T_Y\right)(x, y, \tau), \quad i = 1, 2. \tag{52}$$

Consider the term $[\mathbf{C}_{11}(\theta_c)]_{12}$, where $\mathbf{C}_{11}(\theta_c)$ is given by Eq. 15. Using Eqs. 50, 51 and 52 we have

$$
\begin{aligned}
[\mathbf{C}_{11}(\theta_c)]_{12} &= \int_{t_0}^{t}\int_{\mathbb{R}^2} \frac{1}{\Lambda_1(\tau)q_1\left(x+\frac{d}{2},y\right)+\Lambda_2(\tau)q_2\left(x-\frac{d}{2},y\right)} \\
&\quad\times\left(\Lambda_1(\tau)\frac{\partial q_1\left(x+\frac{d}{2},y\right)}{\partial x}\right)\left(\Lambda_2(\tau)\frac{\partial q_1\left(x+\frac{d}{2},y\right)}{\partial y}\right)dxdyd\tau \\
&= \int_{t_0}^{t}\int_{\mathbb{R}^2} \frac{1}{U^+_{\theta_c}(x,y,\tau)}\frac{\partial W^1_{\theta_c}(x,y,\tau)}{\partial x}\frac{\partial W^1_{\theta_c}(x,y,\tau)}{\partial y}dxdyd\tau \\
&= -\int_{t_0}^{t}\int_{\mathbb{R}^2} \frac{1}{(U^+_{\theta_c}\circ T_Y)(x,y,\tau)} \\
&\quad\times\left(\frac{\partial W^1_{\theta_c}}{\partial x}\circ T_Y\right)(x,y,\tau)\left(\frac{\partial W^1_{\theta_c}}{\partial y}\circ T_Y\right)(x,y,\tau)dxdyd\tau \\
&= -\int_{t_0}^{t}\int_{\mathbb{R}^2}\left(\left(\frac{1}{U^+_{\theta_c}}\frac{\partial W^1_{\theta_c}}{\partial x}\frac{\partial W^1_{\theta_c}}{\partial y}\right)\circ T_Y\right)(x,y,\tau)dxdyd\tau \\
&= \int_{t_0}^{t}\int_{\mathbb{R}^2} \frac{1}{U^+_{\theta_c}(x,y,\tau)}\frac{\partial W^1_{\theta_c}(x,y,\tau)}{\partial x}\frac{\partial W^1_{\theta_c}(x,y,\tau)}{\partial y}dxdyd\tau \\
&= -[\mathbf{C}_{11}(\theta_c)]_{12}, \quad \theta_c\in\Theta_c\backslash\Theta^0_c,
\end{aligned}
$$

where we have used the change of variables theorem in the final step. From the above equation it follows that $[\mathbf{C}_{11}(\theta_c)]_{12}=[\mathbf{C}_{11}(\theta_c)]_{21}=0$, $\theta_c\in\Theta_c\backslash\Theta^0_c$. Similarly, by using Eqs. 50, 51 and 52, we can show that $[\mathbf{C}_{12}(\theta_c)]_{12}=[\mathbf{C}_{12}(\theta_c)]_{21}=0$, and $[\mathbf{C}_{22}(\theta_c)]_{12}=[\mathbf{C}_{22}(\theta_c)]_{21}=0$ for $\theta_c\in\Theta_c\backslash\Theta^0_c$. From this the result follows. □

Lemma 4 *For* $\theta_c=(x_{01},y_{01},x_{02},y_{02})\in\Theta_c$, *let* $\mathbf{K}_{12}(\theta_c)$ *be given by* Eq. 12 *and for* $i=1,2$ *let* $\mathbf{Q}_i$ *be given by* Eq. 20. *Then for* $\theta_c\in\Theta_c$ *and* $i,j=1,2$, *we have*

$$[\mathbf{K}_{12}(\theta_c)]_{ij}\le\sqrt{[\mathbf{Q}_1]_{ii}[\mathbf{Q}_2]_{jj}}<\infty.$$

Proof Define $\Delta_x=x_{02}-x_{01}$ and $\Delta_y=y_{02}-y_{01}$. Applying the Cauchy-Schwarz inequality to the term $[\mathbf{K}_{12}(\theta_c)]_{11}$ and using the fact that $\Lambda_1,\Lambda_2,q_1,q_2\ge 0$, we have for $\theta_c\in\Theta_c$

$$
\begin{aligned}
[\mathbf{K}_{12}(\theta_c)]_{11} &= \int_{t_0}^{t}\int_{\mathbb{R}^2}\frac{\Lambda_1(\tau)\Lambda_2(\tau)}{\Lambda_1(\tau)q_1(x,y)+\Lambda_2(\tau)q_2(x-\Delta_x,y-\Delta_y)} \\
&\quad\times\frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_2(x-\Delta_x,y-\Delta_y)}{\partial x}dxdyd\tau \\
&\le \left(\int_{t_0}^{t}\int_{\mathbb{R}^2}\frac{\Lambda_1(\tau)\Lambda_2(\tau)}{\Lambda_1(\tau)q_1(x,y)+\Lambda_2(\tau)q_2(x-\Delta_x,y-\Delta_y)}\left(\frac{\partial q_1(x,y)}{\partial x}\right)^2dxdyd\tau\right)^{\frac{1}{2}}
\end{aligned}
$$

$$\times \left(\int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{\Lambda_1(\tau)\Lambda_2(\tau)}{\Lambda_1(\tau)q_1(x,y)+\Lambda_2(\tau)q_2(x-\Delta_x,y-\Delta_y)} \left(\frac{\partial q_2(x-\Delta_x,y-\Delta_y)}{\partial x} \right)^2 dxdyd\tau \right)^{\frac{1}{2}}$$

$$\leq \left(\int_{t_0}^{t} \Lambda_2(\tau)d\tau \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^2} \frac{1}{q_1(x,y)} \left(\frac{\partial q_1(x,y)}{\partial x} \right)^2 dxdyd\tau \right)^{\frac{1}{2}}$$

$$\times \left(\int_{t_0}^{t} \Lambda_1(\tau)d\tau \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^2} \frac{1}{q_2(x-\Delta_x,y-\Delta_y)} \left(\frac{\partial q_2(x-\Delta_x,y-\Delta_y)}{\partial x} \right)^2 dxdy \right)^{\frac{1}{2}}$$

$$= \left(\int_{t_0}^{t} \Lambda_1(\tau)d\tau \int_{\mathbb{R}^2} \frac{1}{q_1(x,y)} \left(\frac{\partial q_1(x,y)}{\partial x} \right)^2 dxdy \right)^{\frac{1}{2}}$$

$$\times \left(\int_{t_0}^{t} \Lambda_2(\tau)d\tau \int_{\mathbb{R}^2} \frac{1}{q_2(x,y)} \left(\frac{\partial q_2(x,y)}{\partial x} \right)^2 dxdy \right)^{\frac{1}{2}} = \sqrt{[\mathbf{Q}_1]_{11}[\mathbf{Q}_2]_{22}} < \infty,$$

where we have used the shift invariant property of Lebesgue integrals in the penultimate step, and we have used the properties of image functions (see Definition 6) in the last step. Similarly, we can prove the other results. □

Proof of Theorem 3

Proof Consider the term $\mathbf{K}_{11}(\theta_c)$ given in Eq. 12. By definition, the integral expression of $\mathbf{K}_{11}(\theta_c)$ is measurable for every $\theta_c \in \Theta_c$. Define $\Delta_x := x_{02} - x_{01}$ and $\Delta_y := y_{02} - y_{01}$. Using the shift invariant property of Lebesgue integrals, and the fact that $q_i(x,y) \geq 0$ and $\Lambda_i(\tau) \geq 0$ for $i = 1, 2$, $(x, y) \in \mathbb{R}^2$ and $\tau \geq t_0$, we have for $\theta_c \in \Theta_c$

$$\mathbf{K}_{11}(\theta_c) := \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{\Lambda_1^2(\tau)}{\Lambda_1(\tau)q_1(x,y) + \Lambda_2(\tau)q_2(x-\Delta_x,y-\Delta_y)}$$

$$\times \begin{pmatrix} \left(\frac{\partial q_1(x,y)}{\partial x}\right)^2 & \frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_1(x,y)}{\partial y} \\ \frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_1(x,y)}{\partial y} & \left(\frac{\partial q_1(x,y)}{\partial y}\right)^2 \end{pmatrix} dxdyd\tau$$

$$\leq \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{\Lambda_1^2(\tau)}{\Lambda_1(\tau)q_1(x,y)} \begin{pmatrix} \left(\frac{\partial q_1(x,y)}{\partial x}\right)^2 & \frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_1(x,y)}{\partial y} \\ \frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_1(x,y)}{\partial y} & \left(\frac{\partial q_1(x,y)}{\partial y}\right)^2 \end{pmatrix} dxdyd\tau = \mathbf{Q}_1. \quad (53)$$

By definition of the image function (see Definition 6), we have for $\zeta_1 = x$ and $\zeta_2 = y$, $\int_{\mathbb{R}^2} \frac{1}{q_1(x,y)} \frac{\partial q_1(x,y)}{\partial \zeta_i} \frac{\partial q_1(x,y)}{\partial \zeta_j} dxdy < \infty$ for $i, j = 1, 2$. This implies that $\mathbf{K}_{11}(\theta_c)$ is dominated by the expression given in Eq. 53 for every $\theta_c \in \Theta_c$. By definition of the image

function (see Definition 6), $q_1(x, y)$ and $\frac{\partial q_1(x,y)}{\partial x}$ are continuous for every $x \in \mathbb{R}$. Hence the integrand of $\mathbf{K}_{11}(\theta_c)$ is continuous for every $x \in \mathbb{R}$. Hence by using the Theorem on changing integration and limits for Lebesgue integrals (see Apostol 1974, p. 281), we have

$$\begin{aligned}\lim_{x_{02}\to\infty}\mathbf{K}_{11}(\theta_c) &= \lim_{x_{02}\to\infty}\int_{t_0}^{t}\int_{\mathbb{R}^2}\frac{\Lambda_1^2(\tau)}{\Lambda_1(\tau)q_1(x,y)+\Lambda_2(\tau)q_2(x-\Delta_x,y-\Delta_y)}\\ &\quad\times\begin{pmatrix}\left(\frac{\partial q_1(x,y)}{\partial x}\right)^2 & \frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_1(x,y)}{\partial y}\\ \frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_1(x,y)}{\partial y} & \left(\frac{\partial q_1(x,y)}{\partial y}\right)^2\end{pmatrix}dxdyd\tau\\ &=\int_{t_0}^{t}\int_{\mathbb{R}^2}\lim_{x_{02}\to\infty}\frac{\Lambda_1^2(\tau)}{\Lambda_1(\tau)q_1(x,y)+\Lambda_2(\tau)q_2(x-\Delta_x,y-\Delta_y)}\\ &\quad\times\begin{pmatrix}\left(\frac{\partial q_1(x,y)}{\partial x}\right)^2 & \frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_1(x,y)}{\partial y}\\ \frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_1(x,y)}{\partial y} & \left(\frac{\partial q_1(x,y)}{\partial y}\right)^2\end{pmatrix}dxdyd\tau\\ &=\int_{t_0}^{t}\Lambda_1(\tau)d\tau\int_{\mathbb{R}^2}\frac{1}{q_1(x,y)}\begin{pmatrix}\left(\frac{\partial q_1(x,y)}{\partial x}\right)^2 & \frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_1(x,y)}{\partial y}\\ \frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_1(x,y)}{\partial y} & \left(\frac{\partial q_1(x,y)}{\partial y}\right)^2\end{pmatrix}dxdy=\mathbf{Q}_1,\end{aligned}$$

where we have used assumption **A1** in the next to last step. Similarly, we can show that $\lim_{x_{02}\to\infty}\mathbf{K}_{22}(\theta_c) = \mathbf{Q}_2$. For the term $\mathbf{K}_{12}(\theta_c)$, by definition, the integrand is measurable. Further by definition of the image function, the integrand of $\mathbf{K}_{12}(\theta_c)$ is continuous for every $x \in \mathbb{R}$. From Lemma 4 we know that the entries of $\mathbf{K}_{12}(\theta_c)$ are dominated by integral expressions that are independent of $\theta_c \in \Theta_c$ and are bounded. Hence using the above results pertaining to $\mathbf{K}_{12}(\theta_c)$ and assumptions **A1** and **A2**, we apply the Theorem on changing integration and limits for Lebesgue integrals (see Apostol 1974, p. 281) to obtain

$$\begin{aligned}\lim_{x_{02}\to\infty}\mathbf{K}_{12}(\theta_c) &= \int_{t_0}^{t}\int_{\mathbb{R}^2}\lim_{x_{02}\to\infty}\frac{\Lambda_1(\tau)\Lambda_2(\tau)}{\Lambda_1(\tau)q_1(x,y)+\Lambda_2(\tau)q_2(x-\Delta_x,y-\Delta_y)}\\ &\quad\times\begin{pmatrix}\frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_2(x-\Delta_x,y-\Delta_y)}{\partial x} & \frac{\partial q_1(x,y)}{\partial x}\frac{\partial q_2(x-\Delta_x,y-\Delta_y)}{\partial y}\\ \frac{\partial q_1(x,y)}{\partial y}\frac{\partial q_2(x-\Delta_x,y-\Delta_y)}{\partial x} & \frac{\partial q_1(x,y)}{\partial y}\frac{\partial q_2(x-\Delta_x,y-\Delta_y)}{\partial y}\end{pmatrix}dxdyd\tau\\ &=\int_{t_0}^{t}\int_{\mathbb{R}^2}\frac{\Lambda_1(\tau)\Lambda_2(\tau)}{\Lambda_1(\tau)q_1(x,y)+0}\begin{pmatrix}\frac{\partial q_1(x,y)}{\partial x}0 & \frac{\partial q_1(x,y)}{\partial x}0\\ \frac{\partial q_1(x,y)}{\partial y}0 & \frac{\partial q_1(x,y)}{\partial y}0\end{pmatrix}dxdyd\tau=0.\end{aligned}$$

□

Proof of Theorem 4

Proof 1. The image detection processes $\mathcal{G}_1$ and $\mathcal{G}_2$, which describe the first and second images, respectively, are assumed to be statistically independent of each other. Hence the general expression for the Fisher information matrix can be written as

$$\mathbf{S}_{sim,sp}(\theta_c) = \mathbf{S}_{sim,sp,1}(\theta_c) + \mathbf{S}_{sim,sp,2}(\theta_c), \quad \theta_c \in \Theta_c,$$

where $\mathbf{S}_{sim,sp,1}$ and $\mathbf{S}_{sim,sp,2}$ denote the Fisher information matrices corresponding to the image detection processes $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively. In the present case, we assume without loss of generality that (x_{01}, y_{01}) to be the location coordinates that is determined from the first image. Then it immediately follows that $\mathbf{S}_{sim,sp,1}(\theta_c) = \mathbf{Q}_1$ for $\theta_c \in \Theta_c$, where $\mathbf{Q}_1$ denotes the Fisher information matrix for the localization accuracy problem corresponding to object 1 and is given by Eq. 20.

To derive an expression for $\mathbf{S}_{sim,sp,2}(\theta_c)$, we make use of the fact that for the second image the location coordinates (x_{01}, y_{01}) of object 1 can be assumed to be known a priori, since it is already determined from the first image. Hence for the second image only the location coordinates (x_{02}, y_{02}) of the second object are the unknown parameters. Hence from this it immediately follows that the expression for $\mathbf{S}_{sim,sp,2}(\theta_c)$ will be identical to $\mathbf{K}_{22}(\theta_c)$ which is a component of the Fisher information matrix $\mathbf{S}_{sim}(\theta_c)$ for the problem of estimating θ_c when the location coordinates of both objects are unknown (Theorem 2).

2. To show that $\mathbf{S}_{sim,sp}(\theta_c)$ is invertible, we require that $\mathbf{Q}_1^{-1}$ and $\mathbf{K}_{22}^{-1}(\theta_c)$ exist for every $\theta_c \in \Theta_c$. We prove the result by contradiction. Define $\Delta_x := x_{01} - x_{02}$ and $\Delta_y = y_{01} - y_{02}$. For $\theta_c \in \Theta_c$, consider the term

$$\begin{aligned}
\mathbf{K}_{22}(\theta_c) &= \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{\Lambda_2^2(\tau)}{\Lambda_1(\tau) q_1(x - x_{01}, y - y_{01}) + \Lambda_2(\tau) q_2(x - x_{02}, y - y_{02})} \\
&\quad \times \begin{pmatrix} \left(\frac{\partial q_2(x-x_{02}, y-y_{02})}{\partial x}\right)^2 & \frac{\partial q_2(x-x_{02}, y-y_{02})}{\partial x} \frac{\partial q_2(x-x_{02}, y-y_{02})}{\partial y} \\ \frac{\partial q_2(x-x_{02}, y-y_{02})}{\partial x} \frac{\partial q_2(x-x_{02}, y-y_{02})}{\partial y} & \left(\frac{\partial q_2(x-x_{02}, y-y_{02})}{\partial y}\right)^2 \end{pmatrix} dx dy d\tau \\
&= \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{\Lambda_2^2(\tau)}{\Lambda_1(\tau) q_1(x - \Delta_x, y - \Delta_y) + \Lambda_2(\tau) q_2(x, y)} \\
&\quad \times \begin{pmatrix} \left(\frac{\partial q_2(x,y)}{\partial x}\right)^2 & \frac{\partial q_2(x,y)}{\partial x} \frac{\partial q_2(x,y)}{\partial y} \\ \frac{\partial q_2(x,y)}{\partial x} \frac{\partial q_2(x,y)}{\partial y} & \left(\frac{\partial q_2(x,y)}{\partial y}\right)^2 \end{pmatrix} dx dy d\tau \\
&= \int_{t_0}^{t} \int_{\mathbb{R}^2} h_{\theta_c}(x, y, \tau) \begin{pmatrix} \left(\frac{\partial q_2(x,y)}{\partial x}\right)^2 & \frac{\partial q_2(x,y)}{\partial x} \frac{\partial q_2(x,y)}{\partial y} \\ \frac{\partial q_2(x,y)}{\partial x} \frac{\partial q_2(x,y)}{\partial y} & \left(\frac{\partial q_2(x,y)}{\partial y}\right)^2 \end{pmatrix} dx dy d\tau, \qquad (54)
\end{aligned}$$

where for $\theta_c \in \Theta_c$,

$$h_{\theta_c}(x, y, \tau) := \frac{\Lambda_2^2(\tau)}{\Lambda_1(\tau) q_1(x - \Delta_x, y - \Delta_y) + \Lambda_2(\tau) q_2(x, y)}, \quad (x, y) \in \mathbb{R}^2, \quad \tau \geq t_0.$$

Assume that there exists an image function q_2 such that the Fisher information matrix $\mathbf{K}_{22}(\theta_c)$ is singular for $\theta_c \in \Theta_c$. Hence by Eq. 54, it immediately follow that

$$\begin{aligned}\mathrm{Det}[\mathbf{K}_{22}(\theta_c)] &= \int_{t_0}^{t}\int_{\mathbb{R}^2} h_{\theta_c}(x,y,\tau)\left(\frac{\partial q_2(x,y)}{\partial x}\right)^2 dxdy \\ &\quad\times \int_{t_0}^{t}\int_{\mathbb{R}^2} h_{\theta_c}(x,y,\tau)\left(\frac{\partial q_2(x,y)}{\partial y}\right)^2 dxdy \\ &\quad- \Bigg(\underbrace{\int_{t_0}^{t}\int_{\mathbb{R}^2} h_{\theta_c}(x,y,\tau)\frac{\partial q_2(x,y)}{\partial x}\frac{\partial q_2(x,y)}{\partial y}dxdy}_{T_2}\Bigg)^2 = 0, \quad \theta_c \in \Theta_c.\end{aligned}$$

Note that the above expression pertains to the limiting case of equality of the Cauchy-Schwarz inequality applied to the term T_2. Hence by applying the condition for equality, we have for $k \neq 0$

$$\frac{\partial q_2(x,y)}{\partial x} - k\frac{\partial q_2(x,y)}{\partial y} = 0, \quad (x,y) \in \mathbb{R}^2. \tag{55}$$

The above equation is analogous to the classical one-dimensional transport equation whose solutions are given by (Strauss 1992, p. 6-7)

$$q_2(x,y) = F\left(x + \frac{y}{k}\right), \quad (x,y) \in \mathbb{R}^2,$$

where F is defined on $\mathbb{R}$. As q_2 is an image function satisfying the regularity conditions, we know that q_2 is continuous on $\mathbb{R}^2$. Hence it follows that F is also continuous on $\mathbb{R}$. Further, $q_2(x,y) \geq 0$, $(x,y) \in \mathbb{R}^2$ and hence $F(x) \geq 0$, $x \in \mathbb{R}$. This implies that there exists a constant $K > 0$ and a finite interval $\mathcal{I} = (a,b) \subset \mathbb{R}$ such that $F(x) \geq K$, $x \in \mathcal{I}$. Making use of the fact that $\int_{\mathbb{R}^2} q_2(x,y)dxdy = 1$ (since q_2 is an image function) and substituting for q_2 in terms of F, we have

$$\begin{aligned}1 &= \int_{\mathbb{R}^2} q_2(x,y)dxdy = \int_{\mathbb{R}^2} F\left(x+\frac{y}{k}\right)dxdy = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} F\left(x+\frac{y}{k}\right)dx\right)dy \\ &= \int_{\mathbb{R}}\left(\int_{\mathcal{I}} F\left(x+\frac{y}{k}\right)dx + \int_{\mathcal{I}\setminus\mathbb{R}} F(x+\frac{y}{k})dx\right)dy \\ &\geq \int_{\mathbb{R}}\left(\int_{\mathcal{I}} K dx + \int_{\mathcal{I}\setminus\mathbb{R}} F\left(x+\frac{y}{k}\right)dx\right)dy \\ &= K(b-a)\int_{\mathbb{R}} dy + \int_{\mathbb{R}}\left(\int_{\mathcal{I}\setminus\mathbb{R}} F\left(x+\frac{y}{k}\right)dx\right)dy \\ &= \infty,\end{aligned}$$

which is a contradiction. Hence $\mathbf{K}_{22}(\theta_c)$ is invertible for $\theta_c \in \Theta_c$. Similarly we can show that $\mathbf{Q}_1$ is also invertible. From this the result follows. □

References

Apostol, T. M. (1974). *Mathematical analysis*. Boston, USA: Addison Wesley Publishing Company.

Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., & Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science, 313*, 1642–1645.

Born, M., & Wolf, E. (1999). *Principles of optics*. Cambridge, UK: Cambridge University Press.

Chao, J., Ram, S., Abraham, A., Ward, E. S., & Ober, R. J. (2009a). Resolution in three-dimensional microscopy. *Optics Communications, 282*, 1751–1761.

Chao, J., Ram, S., Ward, E. S., & Ober, R. J. (2009b). A 3D resolution measure for optical microscopy. In *IEEE international symposium on biomedical imaging* (pp. 1115–1118).

Chao, J., Ram, S., Ward, E. S., & Ober, R. J. (2009c). A comparative study of high resolution microscopy imaging modalities using a three-dimensional resolution measure. *Optics Express, 17*, 24,377–24,402.

Gordon, M. P., Ha, T., & Selvin, P. R. (2004). Single molecule high resolution imaging with photobleaching. *Proceedings of the National Academy of Sciences USA, 101*, 6462–6465.

Helstrom, C. W. (1964). The detection and resolution of optical signals. *IEEE Transactions on Information Theory, 10*, 275–287.

Hess, S. T., Girirajan, T. P. K., & Mason, M. D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal, 91*, 4258–4272.

Kay, S. M. (1993). *Fundamentals of statistical signal processing*. New Jersey, USA: Prentice Hall PTR.

Lagerholm, B. C., Averett, L., Weinreb, G. E., Jacobson, K., & Thompson, N. L. (2006). Analysis method for measuring submicroscopic distances with blinking quantum dots. *Biophysical Journal, 91*, 3050–3060.

Lidke, K. A., Rieger, B., Jovin, T. M., & Heintzmann, R. (2005). Superresolution by localization of quantum dots using blinking statistics. *Optics Express, 13*, 7052–7062.

Moerner, W. E. (2007). New directions in single-molecule imaging and analysis. *Proceedings of the National Academy of Sciences USA, 104*, 12,596–12,602.

Ober, R. J., Martinez, C., Lai, X., Zhou, J., & Ward, E. S. (2004a). Exocytosis of IgG as mediated by the receptor, FcRn: an analysis at the single molecule level. *Proceedings of the National Academy of Sciences USA, 101*, 11,076–11,081.

Ober, R. J., Ram, S., & Ward, E. S. (2004b). Localization accuracy in single molecule microscopy. *Biophysical Journal, 86*, 1185–1200.

O'Sullivan, J. A., Blahut, R. E., & Snyder, D. L. (1998). Information-theoretic image formation. *IEEE Transactions on Information Theory, 44*, 2094–2123.

Qu, X., Wu, D., Mets, L., & Scherer, N. F. (2004). Nanometer-localized multiple single-molecule fluorescence microscopy. *Proceedings of the National Academy of Sciences USA, 101*, 11,298–11,303.

Ram, S., Ward, E. S., & Ober, R. J. (2006a). Beyond Rayleigh's criterion: A resolution measure with application to single-molecule microscopy. *Proceedings of the National Academy of Sciences USA, 103*, 4457–4462.

Ram, S., Ward, E. S., & Ober, R. J. (2006b). A stochastic analysis of performance limits for optical microscopes. *Multidimensional Systems and Signal Processing, 17*, 27–58.

Rao, C. R. (1965). *Linear statistical inference and its applications*. New York, USA: Wiley.

Rohr, K. (2007). Theoretical limits of localizing 3D landmarks and features. *IEEE Transactions on Biomedical Engineering, 54*, 1613–1620.

Rudin, W. (1987). *Real and complex analysis*. New York, USA: McGraw Hill.

Rust, M. J., Bates, M., & Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods, 4*, 793–795.

Santos, A., & Young, I. T. (2000). Model-based resolution: Applying the theory in quantitative microscopy. *Applied Optics, 39*, 2948–2958.

Shahram, M., & Milanfar, P. (2004). Imaging below the diffraction limit: A statistical analysis. *IEEE Transactions on Image Processing, 13*, 677–689.

Smith, S. T. (2005). Statistical resolution limits and the complexified Cramer-Rao bound. *IEEE Transactions on Signal Processing, 53*, 1597–1609.

Stoica, P., & Marzetta, T. L. (2001). Parameter estimation problems with singular Fisher information matrices. *IEEE Transactions on Signal Processing, 49*, 87–90.

Strauss, W. A. (1992). *Partial differential equations—an introduction*. New York: Wiley.

Van des Bos, A. (2007). *Parameter estimation for scientists and engineers*. New York, USA: Wiley.
Wong, Y., Lin, Z., & Ober, R. J. (2011). Limit of the accuracy of parameter estimation for moving single molecules imaged by fluorescence microscopy. *IEEE Transactions on Signal Processing, 59*, 895–908.
Young, I. T. (1996). Quantitative microscopy. *IEEE Engineering in Medicine and Biology, 15*, 59–66.
Zhang, F. (1999). *Matrix theory*. New York, USA: Springer.

Author Biographies

Sripad Ram received the B.Sc. degree in applied sciences from PSG College of Technology, Coimbatore, India, in 1999 and the M.Sc. degree in physics from Indian Institute of Technology, Chennai, in 2001. In 2007, he received the Ph.D. degree in Biomedical Engineering from the Joint Biomedical engineering graduate program at the University of Texas at Arlington/the University of Texas Southwestern Medical Center, Dallas. From 2008 to 2011, he held the National Multiple Sclerosis Society postdoctoral fellowship at the University of Texas Southwestern Medical Center, Dallas. He is currently carrying out research in the Department of Immunology, University of Texas Southwestern Medical Center, Dallas. His research interests include statistical image processing of microscopy data, the development of imaging approaches for 3D tracking of single molecules and the study of protein trafficking at cellular barriers.

E. Sally Ward received the Ph.D. degree from the Department of Biochemistry, Cambridge University, Cambridge, U.K., in 1985. From 1985 to 1987, she was a Research Fellow at Gonville and Caius College while working at the Department of Biochemistry, Cambridge University. From 1988 to 1990, she held the Stanley Elmore Senior Research Fellowship at Sidney Sussex College and carried out research at the MRC Laboratory of Molecular Biology, Cambridge. In 1990, she joined the University of Texas Southwestern Medical Center, Dallas, as an Assistant Professor. Since 2002, she has been a Professor in the Department of Immunology at the same institution, and currently holds the Paul and Betty Meek—FINA Professorship in Molecular Immunology. Her research interests include antibody engineering, molecular mechanisms that lead to autoimmune disease, questions related to the in vivo dynamics of antibodies, and the use of microscopy techniques for the study of antibody trafficking in cells.

Raimund J. Ober received the Ph.D. degree in engineering from Cambridge University, Cambridge, U.K., in 1987. From 1987 to 1990, he was a Research Fellow at Girton College and the Engineering Department, Cambridge University. In 1990, he joined the University of Texas at Dallas, Richardson, where he is currently a Professor with the Department of Electrical Engineering. He is also Adjunct Professor at the University of Texas Southwestern Medical Center, Dallas. He is an Associate Editor of *Mathematics of Control, Signals, and Systems*, and a past Associate Editor of *Multidimensional Systems and Signal Processing*, *IEEE Transactions on Circuits and Systems*, and *Systems and Control Letters*. His research interests include the development of microscopy techniques for cellular investigations, in particular at the single molecule level, the study of cellular trafficking pathways using microscopy investigations, and signal/image processing of bioengineering data.